

SEARCH REQUEST FORM

Requestor's

Name: _____

Serial

Number: _____

Date: _____

Phone: _____

Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 5/15/02Searcher: P. Schreiber 308-4292Terminal time: ~~308-4292~~ 8Elapsed time: 9

CPU time: _____

Total time: _____

Number of Searches: _____

Number of Databases: 13

Search Site

____ STIC

☒ CM-1 46A03

____ Pre-S

Type of Search

1 N.A. Sequence1 A.A. Sequence

____ Structure

____ Bibliographic

Vendors

____ IG

____ STN

____ Dialog

____ APS

____ Geninfo

____ SDC

____ DARC/Questel

☒ Other CompuServe

THIS PAGE BLANK (USPTO)

WEST[Help](#)[Logout](#)[Interrupt](#)[Main Menu](#)[Search Form](#)[Posting Counts](#)[Show S Numbers](#)[Edit S Numbers](#)[Preferences](#)[Cases](#)**Search Results -**

Terms	Documents
l5 and l3	0

US Patents Full-Text Database
 US Pre-Grant Publication Full-Text Database
 JPO Abstracts Database
 EPO Abstracts Database
 Derwent World Patents Index
Database: IBM Technical Disclosure Bulletins

Search: L9

[Refine Search](#)[Recall Text](#)[Clear](#)**Search History**
DATE: Wednesday, May 15, 2002
[Printable Copy](#)
[Create Case](#)
Set Name Query
 side by side

Hit Count Set Name
 result set

DB=USPT,PGPB,JPAB,EPAB,DWPI; PLUR=YES; OP=OR

<u>L9</u>	l5 and l3	0	<u>L9</u>
<u>L8</u>	l4 and l3	0	<u>L8</u>
<u>L7</u>	L2 and l3	1	<u>L7</u>
<u>L6</u>	l1 and l3	14	<u>L6</u>
<u>L5</u>	l1 same ny\$	60	<u>L5</u>
<u>L4</u>	l1 same nek\$	5	<u>L4</u>
<u>L3</u>	walke.in. or turner.in. or zambrowicz.in.	9697	<u>L3</u>
<u>L2</u>	L1 same (serin\$ or threoni\$)	262	<u>L2</u>
<u>L1</u>	huma\$ same kinas\$ same (recombin\$ or isolat\$ or clon\$)	2837	<u>L1</u>

THIS PAGE BLANK (USPTO)

WEST[Generate Collection](#)[Print](#)

Search Results - Record(s) 1 through 1 of 1 returned.

☐ 1. Document ID: WO 200142435 A2, AU 200125767 A

L7: Entry 1 of 1

File: DWPI

Jun 14, 2001

DERWENT-ACC-NO: 2001-381667

DERWENT-WEEK: 200140

COPYRIGHT 2002 DERWENT INFORMATION LTD

TITLE: Novel isolated human kinase polynucleotide that shares structural similarity with animal kinases including calcium/calmodulin-dependent protein kinases and serine/threonine protein kinases, useful in therapeutics

INVENTOR: ABUIN, A; DONOHO, G ; FRIEDRICH, G ; SANDS, A T ; SCOVILLE, J ; TURNER, C A ; ZAMBROWICZ, B

PRIORITY-DATA: 1999US-169428P (December 7, 1999)

PATENT-FAMILY:

PUB-NO	PUB-DATE	LANGUAGE	PAGES	MAIN-IPC
WO 200142435 A2	June 14, 2001	E	031	C12N009/00
AU 200125767 A	June 18, 2001		000	C12N009/00

INT-CL (IPC): C12 N 9/00

Full	Title	CIT-1	REV-1	CLS-1	REF-1	SEQ-1	ATT-1
CAW-1							

[Generate Collection](#)[Print](#)

Terms	Documents
L2 and I3	1

Display Format: -

[Change Format](#)[Previous Page](#)[Next Page](#)

THIS PAGE BLANK (USPTO)

66764

Schreiber, David

From: Ramirez, Delia
Sent: Tuesday, May 14, 2002 10:19 AM
To: Schreiber, David
Subject: case 09/783,320

Hi David,

I was wondering if you could do a search for me. I would like to get a standard search of seq id 4 in the nucleic and protein databases (commercial and interference).

Thank you,

Delia

Delia M. Ramirez, Ph.D.
Patent Examiner - Art Unit 1652
USPTO
1911 S. Clark Street, Crystal Mall 1, 10D04, Mail room 10C01
Arlington, VA 22202
(703) 306-0288
delia.ramirez@uspto.gov

THIS PAGE BLANK (USPTO)

=> d his

(FILE 'HOME' ENTERED AT 15:57:10 ON 15 MAY 2002)

INDEX 'ADISALERTS, ADISINSIGHT, ADISNEWS, AGRICOLA, ANABSTR, AQUASCI,
BIOBUSINESS, BIOCOMMERCE, BIOSIS, BIOTECHABS, BIOTECHDS, BIOTECHNO, CABA,
CANCERLIT, CAPLUS, CEABA-VTB, CEN, CIN, CONFSCI, CROPB, CROPU, DDFB,
DDFU, DGENE, DRUGB, DRUGLAUNCH, DRUGMONOG2, ...' ENTERED AT 15:57:19 ON
15 MAY 2002

SEA HUMAN? AND KINAS?

475 FILE ADISALERTS
295 FILE ADISINSIGHT
54 FILE ADISNEWS
297 FILE AGRICOLA
96 FILE ANABSTR
143 FILE AQUASCI
407 FILE BIOBUSINESS
141 FILE BIOCOMMERCE
76512 FILE BIOSIS
1664 FILE BIOTECHABS
1664 FILE BIOTECHDS
32983 FILE BIOTECHNO
1347 FILE CABA
34379 FILE CANCERLIT
46764 FILE CAPLUS
112 FILE CEABA-VTB
33 FILE CEN
77 FILE CIN
569 FILE CONFSCI
27 FILE CROPU
394 FILE DDFB
7268 FILE DDFU
52947 FILE DGENE
394 FILE DRUGB
82 FILE DRUGNL
8737 FILE DRUGU
123 FILE DRUGUPDATES
481 FILE EMBAL
61823 FILE EMBASE
18573 FILE ESBIODBASE
28 FILE FROSTI
29 FILE FSTA
37193 FILE GENBANK
15 FILE HEALSAFE
1105 FILE IFIPAT
5593 FILE JICST-EPLUS
71 FILE KOSMET
13635 FILE LIFESCI
8 FILE MEDICONF
79649 FILE MEDLINE
124 FILE NIOSHTIC
362 FILE NTIS
18 FILE OCEAN
20701 FILE PASCAL
111 FILE PHAR
1 FILE PHIC
157 FILE PHIN
1356 FILE PROMT
49394 FILE SCISEARCH
2 FILE SYNTHLINE
30154 FILE TOXCENTER
19045 FILE USPATFULL
55 FILE USPAT2
1770 FILE WPIDS
1770 FILE WPINDEX
34 FILE IPA
186 FILE NAPRALERT
1843 FILE NLDB

THIS PAGE BLANK (USPTO)

L1 QUE HUMAN? AND KINAS?

FILE 'MEDLINE, BIOSIS, EMBASE, DGENE, SCISEARCH, CAPLUS, GENBANK,
CANCERLIT, BIOTECHNO, TOXCENTER, PASCAL, USPATFULL, ESBIODBASE, LIFESCI,
DRUGU, JICST-EPLUS, NLDB, WPIDS' ENTERED AT 16:00:01 ON 15 MAY 2002

L2 296040 S HUMAN? (S) KINAS?
L3 73739 S L2 (S) (RECOMBIN? OR ISOLAT? OR CLON?)
L4 7879 S L3 (S) (SERIN? OR THREON?)
L5 192 S L3 (S) NEK?
L6 286 S L3 (S) NY?
L7 111 DUP REM L5 (81 DUPLICATES REMOVED)
L8 234 DUP REM L6 (52 DUPLICATES REMOVED)
E WALKER
L9 241 S E3
E TURNER/AU
L10 1510 S E4
E TURNER
L11 64367 S E3
E WALKER/AU
L12 0 S E3
E ZAMBROWICZ/AU
L13 1088 S E4
L14 42 S L3 AND (L12 OR L10 OR L13)
L15 8 S L4 AND (L12 OR L10 OR L13)
L16 0 S L7 AND (L12 OR L10 OR L13)
L17 0 S L8 AND (L12 OR L10 OR L13)
L18 30 DUP REM L14 (12 DUPLICATES REMOVED)

THIS PAGE BLANK (USPTO)

Welcome to STN International! Enter x:x

LOGINID:sssptal652dmr

PASSWORD:

TERMINAL (ENTER 1, 2, 3, OR ?):2

* * * * * Welcome to STN International * * * * *

NEWS 1 Web Page URLs for STN Seminar Schedule - N. America
NEWS 2 Jan 25 BLAST(R) searching in REGISTRY available in STN on the Web
NEWS 3 Jan 29 FSTA has been reloaded and moves to weekly updates
NEWS 4 Feb 01 DKILIT now produced by FIZ Karlsruhe and has a new update frequency
NEWS 5 Feb 19 Access via Tymnet and SprintNet Eliminated Effective 3/31/02
NEWS 6 Mar 08 Gene Names now available in BIOSIS
NEWS 7 Mar 22 TOXLIT no longer available
NEWS 8 Mar 22 TRCTHERMO no longer available
NEWS 9 Mar 28 US Provisional Priorities searched with P in CA/CAPLUS and USPATFULL
NEWS 10 Mar 28 LIPINSKI/CALC added for property searching in REGISTRY
NEWS 11 Apr 02 PAPERCHEM no longer available on STN. Use PAPERCHEM2 instead.
NEWS 12 Apr 08 "Ask CAS" for self-help around the clock
NEWS 13 Apr 09 BEILSTEIN: Reload and Implementation of a New Subject Area
NEWS 14 Apr 09 ZDB will be removed from STN
NEWS 15 Apr 19 US Patent Applications available in IFICDB, IFIPAT, and IFIUDB
NEWS 16 Apr 22 Records from IP.com available in CAPLUS, HCAPLUS, and ZCAPLUS
NEWS 17 Apr 22 BIOSIS Gene Names now available in TOXCENTER
NEWS 18 Apr 22 Federal Research in Progress (FEDRIP) now available

NEWS EXPRESS February 1 CURRENT WINDOWS VERSION IS V6.0d,
CURRENT MACINTOSH VERSION IS V6.0a(ENG) AND V6.0Ja(JP),
AND CURRENT DISCOVER FILE IS DATED 05 FEBRUARY 2002
NEWS HOURS STN Operating Hours Plus Help Desk Availability
NEWS INTER General Internet Information
NEWS LOGIN Welcome Banner and News Items
NEWS PHONE Direct Dial and Telecommunication Network Access to STN
NEWS WWW CAS World Wide Web Site (general information)

Enter NEWS followed by the item number or name to see news on that specific topic.

All use of STN is subject to the provisions of the STN Customer agreement. Please note that this agreement limits use to scientific research. Use for software development or design or implementation of commercial gateways or other similar uses is prohibited and may result in loss of user privileges and other penalties.

* * * * * STN Columbus * * * * *

FILE 'HOME' ENTERED AT 15:57:10 ON 15 MAY 2002

=> index bioscience medicine

FILE 'DRUGMONOG' ACCESS NOT AUTHORIZED

COST IN U.S. DOLLARS

SINCE FILE	TOTAL
ENTRY	SESSION
0.21	0.21

FULL ESTIMATED COST

THIS PAGE BLANK (USPTO)

INDEX 'ADISALERTS, ADISINSIGHT, ADISNEWS, AGRICOLA, ANABSTR, AQUASCI,
BIOBUSINESS, BIOCOMMERCE, BIOSIS, BIOTECHABS, BIOTECHDS, BIOTECHNO, CABA,
CANCERLIT, CAPLUS, CEABA-VTB, CEN, CIN, CONFSCI, CROPB, CROPU, DDFB,
DDFU, DGENE, DRUGB, DRUGLAUNCH, DRUGMONOG2, ...'

ENTERED AT 15:57:19 ON 15 MAY 2002

63 FILES IN THE FILE LIST IN STNINDEX

Enter SET DETAIL ON to see search term postings or to view
search error messages that display as 0* with SET DETAIL OFF.

=> s human? and kinas?

475	FILE ADISALERTS
295	FILE ADISINSIGHT
54	FILE ADISNEWS
297	FILE AGRICOLA
96	FILE ANABSTR
143	FILE AQUASCI
407	FILE BIOBUSINESS
141	FILE BIOCOMMERCE
76512	FILE BIOSIS
1664	FILE BIOTECHABS
1664	FILE BIOTECHDS
32983	FILE BIOTECHNO
1347	FILE CABA
34379	FILE CANCERLIT
46764	FILE CAPLUS
112	FILE CEABA-VTB
33	FILE CEN

17 FILES SEARCHED...

77	FILE CIN
569	FILE CONFSCI
27	FILE CROPU
394	FILE DDFB
7268	FILE DDFU
52947	FILE DGENE
394	FILE DRUGB
82	FILE DRUGNL
8737	FILE DRUGU
123	FILE DRUGUPDATES
481	FILE EMBAL
61823	FILE EMBASE

32 FILES SEARCHED...

18573	FILE ES BIOBASE
28	FILE FROSTI
29	FILE FSTA
37193	FILE GENBANK
15	FILE HEALSAFE
1105	FILE IFIPAT
5593	FILE JICST-EPLUS
71	FILE KOSMET
13635	FILE LIFESCI
8	FILE MEDICONF
79649	FILE MEDLINE
124	FILE NIOSHTIC
362	FILE NTIS
18	FILE OCEAN
20701	FILE PASCAL

49 FILES SEARCHED...

111	FILE PHAR
1	FILE PHIC
157	FILE PHIN

THIS PAGE BLANK (USPTO)

1356	FILE PROMT
49394	FILE SCISEARCH
2	FILE SYNTHLINE
30154	FILE TOXCENTER
19045	FILE USPATFULL
55	FILE USPAT2
1770	FILE WPIDS
1770	FILE WPINDEX
34	FILE IPA
186	FILE NAPRALERT
1843	FILE NLDB

58 FILES HAVE ONE OR MORE ANSWERS, 63 FILES SEARCHED IN STNINDEX

L1 QUE HUMAN? AND KINAS?

=> d rank

F1	79649	MEDLINE
F2	76512	BIOSIS
F3	61823	EMBASE
F4	52947	DGENE
F5	49394	SCISEARCH
F6	46764	CAPLUS
F7	37193	GENBANK
F8	34379	CANCERLIT
F9	32983	BIOTECHNO
F10	30154	TOXCENTER
F11	20701	PASCAL
F12	19045	USPATFULL
F13	18573	ESBIOBASE
F14	13635	LIFESCI
F15	8737	DRUGU
F16	7268	DDFU
F17	5593	JICST-EPLUS
F18	1843	NLDB
F19	1770	WPIDS
F20	1770	WPINDEX
F21	1664	BIOTECHABS
F22	1664	BIOTECHDS
F23	1356	PROMT
F24	1347	CABA
F25	1105	IFIPAT
F26	569	CONFSCI
F27	481	EMBAL
F28	475	ADISALERTS
F29	407	BIOBUSINESS
F30	394	DDFB
F31	394	DRUGB
F32	362	NTIS
F33	297	AGRICOLA
F34	295	ADISINSIGHT
F35	186	NAPRALERT
F36	157	PHIN
F37	143	AQUASCI
F38	141	BIOCOMMERCE
F39	124	NIOSHTIC
F40	123	DRUGUPDATES
F41	112	CEABA-VTB
F42	111	PHAR
F43	96	ANABSTR
F44	82	DRUGNL
F45	77	CIN
F46	71	KOSMET

THIS PAGE BLANK (USPTO)

F47	55	USPAT2
F48	54	ADISNEWS
F49	34	IPA
F50	33	CEN
F51	29	FSTA
F52	28	FROSTI
F53	27	CROPU
F54	18	OCEAN
F55	15	HEALSAFE
F56	8	MEDICONF
F57	2	SYNTHLINE
F58	1	PHIC

=> fil f1-f20

COST IN U.S. DOLLARS

SINCE FILE	TOTAL
ENTRY	SESSION
2.65	2.86

FULL ESTIMATED COST

FILE 'MEDLINE' ENTERED AT 16:00:01 ON 15 MAY 2002

FILE 'BIOSIS' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 BIOLOGICAL ABSTRACTS INC. (R)

FILE 'EMBASE' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 Elsevier Science B.V. All rights reserved.

FILE 'DGENE' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 THOMSON DERWENT

FILE 'SCISEARCH' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 Institute for Scientific Information (ISI) (R)

FILE 'CAPLUS' ENTERED AT 16:00:01 ON 15 MAY 2002

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'GENBANK' ENTERED AT 16:00:01 ON 15 MAY 2002

FILE 'CANCERLIT' ENTERED AT 16:00:01 ON 15 MAY 2002

FILE 'BIOTECHNO' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 Elsevier Science B.V., Amsterdam. All rights reserved.

FILE 'TOXCENTER' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 ACS

FILE 'PASCAL' ENTERED AT 16:00:01 ON 15 MAY 2002

Any reproduction or dissemination in part or in full,
by means of any process and on any support whatsoever
is prohibited without the prior written agreement of INIST-CNRS.

COPYRIGHT (C) 2002 INIST-CNRS. All rights reserved.

FILE 'USPATFULL' ENTERED AT 16:00:01 ON 15 MAY 2002

CA INDEXING COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'ESBIOBASE' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 Elsevier Science B.V., Amsterdam. All rights reserved.

FILE 'LIFESCI' ENTERED AT 16:00:01 ON 15 MAY 2002

COPYRIGHT (C) 2002 Cambridge Scientific Abstracts (CSA)

THIS PAGE BLANK (USPTO)

FILE 'DRUGU' ENTERED AT 16:00:01 ON 15 MAY 2002
COPYRIGHT (C) 2002 THOMSON DERWENT

FILE 'DDFU' ACCESS NOT AUTHORIZED

FILE 'JICST-EPLUS' ENTERED AT 16:00:01 ON 15 MAY 2002
COPYRIGHT (C) 2002 Japan Science and Technology Corporation (JST)

FILE 'NLDB' ENTERED AT 16:00:01 ON 15 MAY 2002
COPYRIGHT (C) 2002 Gale Group. All rights reserved.

FILE 'WPIDS' ENTERED AT 16:00:01 ON 15 MAY 2002
COPYRIGHT (C) 2002 THOMSON DERWENT

FILE 'WPINDEX' ACCESS NOT AUTHORIZED

=> s human? (s) kinas?

4 FILES SEARCHED...

9 FILES SEARCHED...

11 FILES SEARCHED...

L2 296040 HUMAN? (S) KINAS?

=> s l2 (s) (recombin? or isolat? or clon?)

4 FILES SEARCHED...

7 FILES SEARCHED...

12 FILES SEARCHED...

L3 73739 L2 (S) (RECOMBIN? OR ISOLAT? OR CLON?)

<-----User Break----->

u

SEARCH ENDED BY USER

=> s l3 (s) (serin? or threon?)

L4 7879 L3 (S) (SERIN? OR THREON?)

=> s l3 (s) nek?

L5 192 L3 (S) NEK?

=> s l3 (s) ny?

L6 286 L3 (S) NY?

=> dup rem l5

DUPLICATE IS NOT AVAILABLE IN 'DGENE, GENBANK'.

ANSWERS FROM THESE FILES WILL BE CONSIDERED UNIQUE

PROCESSING COMPLETED FOR L5

L7 111 DUP REM L5 (81 DUPLICATES REMOVED)

=> dup rem l6

DUPLICATE IS NOT AVAILABLE IN 'DGENE, GENBANK'.

ANSWERS FROM THESE FILES WILL BE CONSIDERED UNIQUE

PROCESSING COMPLETED FOR L6

L8 234 DUP REM L6 (52 DUPLICATES REMOVED)

=> e walke

E1 12 WALKDOWNS/BI

E2 7 WALKDR4/BI

E3 241 --> WALKE/BI

E4 1 WALKEAKOSKI/BI

E5 1 WALKEAYS/BI

E6 15746 WALKED/BI

E7 2 WALKEDDED/BI

E8 1 WALKEDFREQUENTLY/BI

THIS PAGE BLANK (USPTO)

E9	1	WALKEDOFF/BI
E10	1	WALKEDOUT/BI
E11	2	WALKEE/BI
E12	1	WALKEFIELD/BI

=> s e3

L9 241 WALKE/BI

=> e turner/au

'AU' IS NOT A VALID EXPAND FIELD CODE FOR FILE 'NLDB'

E1	1	TURNEN R N/AU
E2	1	TURNENS A P/AU
E3	14	--> TURNER/AU
E4	1510	TURNER A/AU
E5	38	TURNER A A/AU
E6	4	TURNER A ALLAN/AU
E7	164	TURNER A B/AU
E8	249	TURNER A C/AU
E9	223	TURNER A D/AU
E10	3	TURNER A D C/AU
E11	1	TURNER A D M S/AU
E12	8	TURNER A DENISE/AU

The indicated field code is not available for EXPAND in this file. To see a list of valid EXPAND field codes, enter HELP SFIELDS at an arrow prompt (=>).

=> s e4

'AU' IS NOT A VALID FIELD CODE

L10 1510 "TURNER A"/AU

=> e turner

E1	1	TURNEPPE/BI
E2	1	TURNEQ/BI
E3	64367	--> TURNER/BI
E4	1	TURNER1/BI
E5	1	TURNER1852/BI
E6	1	TURNER2000/BI
E7	1	TURNER38/BI
E8	471	TURNERA/BI
E9	327	TURNERACEAE/BI
E10	1	TURNERACEAS/BI
E11	1	TURNERACEES/BI
E12	2	TURNERACEOUS/BI

=> s e3

L11 64367 TURNER/BI

=> e walke/au

'AU' IS NOT A VALID EXPAND FIELD CODE FOR FILE 'NLDB'

E1	1	WALKDHUETE W/AU
E2	1	WALKDORFF H R/AU
E3	0	--> WALKE/AU
E4	3	WALKE A/AU
E5	1	WALKE A E/AU
E6	3	WALKE A L/AU
E7	1	WALKE ALISON B/AU
E8	2	WALKE ALLEN L/AU
E9	50	WALKE B/AU
E10	3	WALKE B H/AU
E11	1	WALKE BERNHARD H/AU
E12	1	WALKE BERTHOLD/AU

The indicated field code is not available for EXPAND in this file. To see a list of valid EXPAND field codes, enter HELP

THIS PAGE BLANK (USPTO)

SFIELDS at an arrow prompt (=>).

=> s e3

'AU' IS NOT A VALID FIELD CODE

L12 0 WALKER/AU

=> e zambrowicz/au

'AU' IS NOT A VALID EXPAND FIELD CODE FOR FILE 'NLDB'

E1	1	ZAMBROWIC B P/AU
E2	1	ZAMBROWICA E B/AU
E3	0 -->	ZAMBROWICZ/AU
E4	1088	ZAMBROWICZ B/AU
E5	4	ZAMBROWICZ B E/AU
E6	93	ZAMBROWICZ B P/AU
E7	2	ZAMBROWICZ BRIAN P/AU
E8	89	ZAMBROWICZ BRIAN/AU
E9	30	ZAMBROWICZ BRIAN P/AU
E10	1	ZAMBROWICZ BRIAN PETER/AU
E11	1	ZAMBROWICZ BRYGIDA E/AU
E12	2	ZAMBROWICZ E/AU

The indicated field code is not available for EXPAND in this file. To see a list of valid EXPAND field codes, enter HELP SFIELDS at an arrow prompt (=>).

=> s e4

'AU' IS NOT A VALID FIELD CODE

L13 1088 "ZAMBROWICZ B"/AU

=> d his

(FILE 'HOME' ENTERED AT 15:57:10 ON 15 MAY 2002)

INDEX 'ADISALERTS, ADISINSIGHT, ADISNEWS, AGRICOLA, ANABSTR, AQUASCI, BIOBUSINESS, BIOCOMMERCE, BIOSIS, BIOTECHABS, BIOTECHDS, BIOTECHNO, CABA, CANCERLIT, CAPLUS, CEABA-VTB, CEN, CIN, CONFSCI, CROPB, CROPU, DDFB, DDFU, DGENE, DRUGB, DRUGLAUNCH, DRUGMONOG2, ...' ENTERED AT 15:57:19 ON 15 MAY 2002

SEA HUMAN? AND KINAS?

475	FILE ADISALERTS
295	FILE ADISINSIGHT
54	FILE ADISNEWS
297	FILE AGRICOLA
96	FILE ANABSTR
143	FILE AQUASCI
407	FILE BIOBUSINESS
141	FILE BIOCOMMERCE
76512	FILE BIOSIS
1664	FILE BIOTECHABS
1664	FILE BIOTECHDS
32983	FILE BIOTECHNO
1347	FILE CABA
34379	FILE CANCERLIT
46764	FILE CAPLUS
112	FILE CEABA-VTB
33	FILE CEN
77	FILE CIN
569	FILE CONFSCI
27	FILE CROPU
394	FILE DDFB
7268	FILE DDFU
52947	FILE DGENE
394	FILE DRUGB

THIS PAGE BLANK (USPTO)

82 FILE DRUGNL
 8737 FILE DRUGU
 123 FILE DRUGUPDATES
 481 FILE EMBAL
 61823 FILE EMBASE
 18573 FILE ESBIODBASE
 28 FILE FROSTI
 29 FILE FSTA
 37193 FILE GENBANK
 15 FILE HEALSAFE
 1105 FILE IFIPAT
 5593 FILE JICST-EPLUS
 71 FILE KOSMET
 13635 FILE LIFESCI
 8 FILE MEDICONF
 79649 FILE MEDLINE
 124 FILE NIOSHTIC
 362 FILE NTIS
 18 FILE OCEAN
 20701 FILE PASCAL
 111 FILE PHAR
 1 FILE PHIC
 157 FILE PHIN
 1356 FILE PROMT
 49394 FILE SCISEARCH
 2 FILE SYNTHLINE
 30154 FILE TOXCENTER
 19045 FILE USPATFULL
 55 FILE USPAT2
 1770 FILE WPIDS
 1770 FILE WPINDEX
 34 FILE IPA
 186 FILE NAPRALERT
 1843 FILE NLDB

L1 QUE HUMAN? AND KINAS?

FILE 'MEDLINE, BIOSIS, EMBASE, DGENE, SCISEARCH, CAPLUS, GENBANK,
 CANCERLIT, BIOTECHNO, TOXCENTER, PASCAL, USPATFULL, ESBIODBASE, LIFESCI,
 DRUGU, JICST-EPLUS, NLDB, WPIDS' ENTERED AT 16:00:01 ON 15 MAY 2002

L2 296040 S HUMAN? (S) KINAS?
 L3 73739 S L2 (S) (RECOMBIN? OR ISOLAT? OR CLON?)
 L4 7879 S L3 (S) (SERIN? OR THREON?)
 L5 192 S L3 (S) NEK?
 L6 286 S L3 (S) NY?
 L7 111 DUP REM L5 (81 DUPLICATES REMOVED)
 L8 234 DUP REM L6 (52 DUPLICATES REMOVED)
 E WALKER
 L9 241 S E3
 E TURNER/AU
 L10 1510 S E4
 E TURNER
 L11 64367 S E3
 E WALKER/AU
 L12 0 S E3
 E ZAMBROWICZ/AU
 L13 1088 S E4

=> s l3 and (l12 or l10 or l13)

15 FILES SEARCHED...

L14 42 L3 AND (L12 OR L10 OR L13)

=> s l4 and (L12 OR L10 OR L13)

THIS PAGE BLANK (USPTO)

L15 8 L4 AND (L12 OR L10 OR L13)

=> s 17 and (L12 OR L10 OR L13)

L16 0 L7 AND (L12 OR L10 OR L13)

=> s 18 and (L12 OR L10 OR L13)

L17 0 L8 AND (L12 OR L10 OR L13)

=> d his

(FILE 'HOME' ENTERED AT 15:57:10 ON 15 MAY 2002)

INDEX 'ADISALERTS, ADISINSIGHT, ADISNEWS, AGRICOLA, ANABSTR, AQUASCI, BIOBUSINESS, BIOCOMMERCE, BIOSIS, BIOTECHABS, BIOTECHDS, BIOTECHNO, CABA, CANCERLIT, CAPLUS, CEABA-VTB, CEN, CIN, CONFSCI, CROPB, CROPU, DDFB, DDFU, DGENE, DRUGB, DRUGLAUNCH, DRUGMONOG2, ...' ENTERED AT 15:57:19 ON 15 MAY 2002

SEA HUMAN? AND KINAS?

475 FILE ADISALERTS
295 FILE ADISINSIGHT
54 FILE ADISNEWS
297 FILE AGRICOLA
96 FILE ANABSTR
143 FILE AQUASCI
407 FILE BIOBUSINESS
141 FILE BIOCOMMERCE
76512 FILE BIOSIS
1664 FILE BIOTECHABS
1664 FILE BIOTECHDS
32983 FILE BIOTECHNO
1347 FILE CABA
34379 FILE CANCERLIT
46764 FILE CAPLUS
112 FILE CEABA-VTB
33 FILE CEN
77 FILE CIN
569 FILE CONFSCI
27 FILE CROPU
394 FILE DDFB
7268 FILE DDFU
52947 FILE DGENE
394 FILE DRUGB
82 FILE DRUGNL
8737 FILE DRUGU
123 FILE DRUGUPDATES
481 FILE EMBAL
61823 FILE EMBASE
18573 FILE ESBIODASE
28 FILE FROSTI
29 FILE FSTA
37193 FILE GENBANK
15 FILE HEALSAFE
1105 FILE IFIPAT
5593 FILE JICST-EPLUS
71 FILE KOSMET
13635 FILE LIFESCI
8 FILE MEDICONF
79649 FILE MEDLINE
124 FILE NIOSHTIC
362 FILE NTIS
18 FILE OCEAN
20701 FILE PASCAL

THIS PAGE BLANK (USPTO)

111 FILE PHAR
 1 FILE PHIC
 157 FILE PHIN
 1356 FILE PROMT
 49394 FILE SCISEARCH
 2 FILE SYNTHLINE
 30154 FILE TOXCENTER
 19045 FILE USPATFULL
 55 FILE USPAT2
 1770 FILE WPIDS
 1770 FILE WPINDEX
 34 FILE IPA
 186 FILE NAPRALERT
 1843 FILE NLDB

L1 QUE HUMAN? AND KINAS?

FILE 'MEDLINE, BIOSIS, EMBASE, DGENE, SCISEARCH, CAPLUS, GENBANK,
 CANCERLIT, BIOTECHNO, TOXCENTER, PASCAL, USPATFULL, ESBIODBASE, LIFESCI,
 DRUGU, JICST-EPLUS, NLDB, WPIDS' ENTERED AT 16:00:01 ON 15 MAY 2002

L2 296040 S HUMAN? (S) KINAS?
 L3 73739 S L2 (S) (RECOMBIN? OR ISOLAT? OR CLON?)
 L4 7879 S L3 (S) (SERIN? OR THREON?)
 L5 192 S L3 (S) NEK?
 L6 286 S L3 (S) NY?
 L7 111 DUP REM L5 (81 DUPLICATES REMOVED)
 L8 234 DUP REM L6 (52 DUPLICATES REMOVED)
 E WALKER
 L9 241 S E3
 E TURNER/AU
 L10 1510 S E4
 E TURNER
 L11 64367 S E3
 E WALKER/AU
 L12 0 S E3
 E ZAMBROWICZ/AU
 L13 1088 S E4
 L14 42 S L3 AND (L12 OR L10 OR L13)
 L15 8 S L4 AND (L12 OR L10 OR L13)
 L16 0 S L7 AND (L12 OR L10 OR L13)
 L17 0 S L8 AND (L12 OR L10 OR L13)

=> dup rem l14

DUPLICATE IS NOT AVAILABLE IN 'DGENE, GENBANK'.

ANSWERS FROM THESE FILES WILL BE CONSIDERED UNIQUE

PROCESSING COMPLETED FOR L14

L18 30 DUP REM L14 (12 DUPLICATES REMOVED)

=> d ti l18 1-30

L18 ANSWER 1 OF 30 WPIDS (C) 2002 THOMSON DERWENT

TI Novel **isolated human kinase** polynucleotide
 that shares structural similarity with animal **kinases** including
 calcium/calmodulin-dependent protein **kinases** and
 serine/threonine protein **kinases**, useful in therapeutics.

L18 ANSWER 2 OF 30 WPIDS (C) 2002 THOMSON DERWENT

TI Novel **isolated human kinase** polynucleotide
 useful for screening for drugs effective in treatment of symptomatic or
 phenotypic manifestations of perturbing normal function of novel
human kinase protein in the body.

L18 ANSWER 3 OF 30 WPIDS (C) 2002 THOMSON DERWENT

THIS PAGE BLANK (USPTO)

TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and for
screening for compounds useful for treating mental, biological or medical
diseases.

L18 ANSWER 4 OF 30 WPIDS (C) 2002 THOMSON DERWENT

TI New substituted 1,4-Dihydroindeno(1,2-C)Pyrazoles useful in the treatment
of e.g. cancer, arthritis and atherosclerosis are Tyrosine Kinase
Inhibitors.

L18 ANSWER 5 OF 30 WPIDS (C) 2002 THOMSON DERWENT

TI Constructing normalized eukaryotic cDNA libraries, useful for isolating
and identifying new cDNA.

L18 ANSWER 6 OF 30 MEDLINE DUPLICATE 1

TI Characterization of a multidrug resistant human erythroleukemia cell line
(K562) exhibiting spontaneous resistance to 1-beta-D-
arabinofuranosylcytosine.

L18 ANSWER 7 OF 30 MEDLINE DUPLICATE 2

TI In vitro effects of bryostatin 1 on the metabolism and cytotoxicity of
1-beta-D-arabinofuranosylcytosine in human leukemia cells.

L18 ANSWER 8 OF 30 DGENE (C) 2002 THOMSON DERWENT

TI An isolated nucleic acid molecule encoding a novel human protein useful
as therapeutics and to screen libraries isolate clones and prepare
cloning and sequencing templates -

L18 ANSWER 9 OF 30 DGENE (C) 2002 THOMSON DERWENT

TI Novel **isolated human kinase** polynucleotide
that shares structural similarity with animal **kinases** including
calcium/calmodulin-dependent protein **kinases** and
serine/threonine protein **kinases**, useful in therapeutics -

L18 ANSWER 10 OF 30 DGENE (C) 2002 THOMSON DERWENT

TI Novel **isolated human kinase** polynucleotide
that shares structural similarity with animal **kinases** including
calcium/calmodulin-dependent protein **kinases** and
serine/threonine protein **kinases**, useful in therapeutics -

L18 ANSWER 11 OF 30 DGENE (C) 2002 THOMSON DERWENT

TI Novel **isolated human kinase** polynucleotide
useful for screening for drugs effective in treatment of symptomatic or
phenotypic manifestations of perturbing normal function of novel
human kinase protein in the body -

L18 ANSWER 12 OF 30 DGENE (C) 2002 THOMSON DERWENT

TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 13 OF 30 DGENE (C) 2002 THOMSON DERWENT

TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 14 OF 30 DGENE (C) 2002 THOMSON DERWENT

TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

THIS PAGE BLANK (USPTO)

L18 ANSWER 15 OF 30 DGENE (C) 2002 THOMSON DERWENT
 TI New **isolated human kinase** polynucleotide
 useful for generating antibodies, as reagents in diagnostic assays and
 for screening for compounds useful for treating mental, biological or
 medical diseases -

L18 ANSWER 16 OF 30 DGENE (C) 2002 THOMSON DERWENT
 TI New **isolated human kinase** polynucleotide
 useful for generating antibodies, as reagents in diagnostic assays and
 for screening for compounds useful for treating mental, biological or
 medical diseases -

L18 ANSWER 17 OF 30 DGENE (C) 2002 THOMSON DERWENT
 TI New **isolated human kinase** polynucleotide
 useful for generating antibodies, as reagents in diagnostic assays and
 for screening for compounds useful for treating mental, biological or
 medical diseases -

L18 ANSWER 18 OF 30 DGENE (C) 2002 THOMSON DERWENT
 TI An isolated nucleic acid molecule encoding a novel human protein useful
 as therapeutics and to screen libraries isolate clones and prepare
 cloning and sequencing templates -

L18 ANSWER 19 OF 30 DGENE (C) 2002 THOMSON DERWENT
 TI An isolated nucleic acid molecule encoding a novel human protein useful
 as therapeutics and to screen libraries isolate clones and prepare
 cloning and sequencing templates -

L18 ANSWER 20 OF 30 DGENE (C) 2002 THOMSON DERWENT
 TI Novel **isolated human kinase** polynucleotide
 that shares structural similarity with animal **kinases** including
 calcium/calmodulin-dependent protein **kinases** and
 serine/threonine protein **kinases**, useful in therapeutics -

L18 ANSWER 21 OF 30 DGENE (C) 2002 THOMSON DERWENT
 TI Novel **isolated human kinase** polynucleotide
 that shares structural similarity with animal **kinases** including
 calcium/calmodulin-dependent protein **kinases** and
 serine/threonine protein **kinases**, useful in therapeutics -

L18 ANSWER 22 OF 30 DGENE (C) 2002 THOMSON DERWENT
 TI Novel **isolated human kinase** polynucleotide
 that shares structural similarity with animal **kinases** including
 calcium/calmodulin-dependent protein **kinases** and
 serine/threonine protein **kinases**, useful in therapeutics -

L18 ANSWER 23 OF 30 DGENE (C) 2002 THOMSON DERWENT
 TI Novel **isolated human kinase** polynucleotide
 useful for screening for drugs effective in treatment of symptomatic or
 phenotypic manifestations of perturbing normal function of novel
human kinase protein in the body -

L18 ANSWER 24 OF 30 DGENE (C) 2002 THOMSON DERWENT
 TI New **isolated human kinase** polynucleotide
 useful for generating antibodies, as reagents in diagnostic assays and
 for screening for compounds useful for treating mental, biological or
 medical diseases -

L18 ANSWER 25 OF 30 DGENE (C) 2002 THOMSON DERWENT
 TI New **isolated human kinase** polynucleotide
 useful for generating antibodies, as reagents in diagnostic assays and
 for screening for compounds useful for treating mental, biological or

THIS PAGE BLANK (USPTO)

medical diseases -

L18 ANSWER 26 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 27 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 28 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 29 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

L18 ANSWER 30 OF 30 DGENE (C) 2002 THOMSON DERWENT
TI New **isolated human kinase** polynucleotide
useful for generating antibodies, as reagents in diagnostic assays and
for screening for compounds useful for treating mental, biological or
medical diseases -

=> d 1

L18 ANSWER 1 OF 30 WPIDS (C) 2002 THOMSON DERWENT
AN 2001-381667 [40] WPIDS
DNC C2001-116955
TI Novel **isolated human kinase** polynucleotide
that shares structural similarity with animal **kinases** including
calcium/calmodulin-dependent protein **kinases** and
serine/threonine protein **kinases**, useful in therapeutics.
DC B04 D16
IN ABUIN, A; DONOHO, G; FRIEDRICH, G; SANDS, A T; SCOVILLE, J; TURNER, C A;
ZAMBROWICZ, B
PA (LEXI-N) LEXICON GENETICS INC
CYC 93
PI WO 2001042435 A2 20010614 (200140)* EN 31p C12N009-00
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ
NL OA PT SD SE SL SZ TR TZ UG ZW
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE DK DM
DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC
LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE
SG SI SK SL TJ TM TR TT TZ UA UG UZ VN YU ZA ZW
AU 2001025767 A 20010618 (200161) C12N009-00
ADT WO 2001042435 A2 WO 2000-US33240 20001207; AU 2001025767 A AU 2001-25767
20001207
FDT AU 2001025767 A Based on WO 200142435
PRAI US 1999-169428P 19991207
IC ICM C12N009-00

=> d 4-9

THIS PAGE BLANK (USPTO)

L18 ANSWER 4 OF 30 WPIDS (C) 2002 THOMSON DERWENT
AN 2000-647337 [62] WPIDS
DNC C2000-195868
TI New substituted 1,4-Dihydroindeno(1,2-C)Pyrazoles useful in the treatment of e.g. cancer, arthritis and atherosclerosis are Tyrosine Kinase Inhibitors.
DC B02
IN ARNOLD, L; DOYLE, K; RAFFERTY, P; STEELE, R; **TURNER, A**; WILKINS, D; ARNOLD, L D; DOYLE, K J; STEELE, R W; WILKINS, D J
PA (BADI) BASF AG; (KNOL) KNOLL GMBH
CYC 43
PI WO 2000059901 A1 20001012 (200062)* EN 78p C07D401-12
RW: AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE
W: AU BG BR CA CN CZ HR HU ID IL IN JP KR MX NO NZ PL RU SG SK TR UA
US ZA
AU 2000040378 A 20001023 (200107) C07D401-12
US 6297238 B1 20011002 (200160) A61K031-4162
EP 1165544 A1 20020102 (200209) EN C07D401-12
R: AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
NO 2001004864 A 20011205 (200210) C07D000-00
CZ 2001003580 A3 20020313 (200223) C07D495-04
ADT WO 2000059901 A1 WO 2000-US8192 20000328; AU 2000040378 A AU 2000-40378 20000328; US 6297238 B1 Provisional US 1999-127963P 19990406, CIP of US 2000-541336 20000403, US 2000-689943 20001012; EP 1165544 A1 EP 2000-919742 20000328, WO 2000-US8192 20000328; NO 2001004864 A WO 2000-US8192 20000328, NO 2001-4864 20011005; CZ 2001003580 A3 WO 2000-US8192 20000328, CZ 2001-3580 20000328
FDT AU 2000040378 A Based on WO 200059901; EP 1165544 A1 Based on WO 200059901; CZ 2001003580 A3 Based on WO 200059901
PRAI US 1999-127963P 19990406; US 2000-541336 20000403; US 2000-689943 20001012
IC ICM A61K031-4162; C07D000-00; C07D401-12; C07D495-04
ICS A61K031-416; A61P001-04; A61P009-10; A61P031-04; A61P035-00; A61P043-00; C07D231-54; C07D403-12

L18 ANSWER 5 OF 30 WPIDS (C) 2002 THOMSON DERWENT
AN 2000-224329 [19] WPIDS
DNC C2000-068498
TI Constructing normalized eukaryotic cDNA libraries, useful for isolating and identifying new cDNA.
DC B04 D16
IN FRIEDRICH, G; NEHLS, M; RULEY, H E; SANDS, A T; WATTLER, S;
ZAMBROWICZ, B
PA (LEXI-N) LEXICON GENETICS INC
CYC 87
PI WO 2000009681 A2 20000224 (200019)* EN 47p C12N015-10
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW NL
OA PT SD SE SL SZ UG ZW
W: AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB
GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU
LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR
TT UA UG UZ VN YU ZA ZW
AU 9954703 A 20000306 (200030) C12N015-10
US 6218123 B1 20010417 (200123) C12Q001-68
EP 1105471 A2 20010613 (200134) EN C12N015-10
R: AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT
RO SE SI
ADT WO 2000009681 A2 WO 1999-US17945 19990810; AU 9954703 A AU 1999-54703 19990810; US 6218123 B1 Provisional US 1998-95989P 19980810, US 1999-371257 19990810; EP 1105471 A2 EP 1999-940956 19990810, WO 1999-US17945 19990810
FDT AU 9954703 A Based on WO 200009681; EP 1105471 A2 Based on WO 200009681

THIS PAGE BLANK (USPTO)

PRAI US 1998-95989P 19980810; US 1999-371257 19990810
IC ICM C12N015-10; C12Q001-68
ICS C12N015-11; C12N015-62; C12N015-63

L18 ANSWER 6 OF 30 MEDLINE DUPLICATE 1
AN 95287645 MEDLINE
DN 95287645 PubMed ID: 7769843
TI Characterization of a multidrug resistant human erythroleukemia cell line (K562) exhibiting spontaneous resistance to 1-beta-D-arabinofuranosylcytosine.
AU Grant S; **Turner A**; Nelms P; Yanovich S
CS Department of Medicine, Medical College of Virginia, Richmond 23298, USA.
NC 1R01 CA63753 (NCI)
CA-16059 (NCI)
SO LEUKEMIA, (1995 May) 9 (5) 808-14.
Journal code: LEU; 8704895. ISSN: 0887-6924.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199507
ED Entered STN: 19950713
Last Updated on STN: 19970203
Entered Medline: 19950706

L18 ANSWER 7 OF 30 MEDLINE DUPLICATE 2
AN 91328750 MEDLINE
DN 91328750 PubMed ID: 1867641
TI In vitro effects of bryostatin 1 on the metabolism and cytotoxicity of 1-beta-D-arabinofuranosylcytosine in human leukemia cells.
AU Grant S; Boise L; Westin E; Howe C; Pettit G R; **Turner A**; McCrady C
CS Division of Hematology/Oncology, Medical College of Virginia, Richmond 23298.
NC AICR88B36 (NIAID)
CA04875 (NCI)
R01 CA35601 (NCI)
SO BIOCHEMICAL PHARMACOLOGY, (1991 Jul 25) 42 (4) 853-67.
Journal code: 9Z4; 0101032. ISSN: 0006-2952.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199109
ED Entered STN: 19910929
Last Updated on STN: 19970203
Entered Medline: 19910909

L18 ANSWER 8 OF 30 DGENE (C) 2002 THOMSON DERWENT
AN AAU04636 Protein DGENE
TI An isolated nucleic acid molecule encoding a novel human protein useful as therapeutics and to screen libraries isolate clones and prepare cloning and sequencing templates -
IN Donoho G; Hilbun E; Turner C A; Friedrich G; **Zambrowicz B**; Sands A T
PA (LEXI-N) LEXICON GENETICS INC.
PI WO 2001053493 A2 20010726 33p
AI WO 2001-US2120 20010118
PRAI US 2000-176690 20000118
DT Patent
LA English
OS 2001-442260 [47]

THIS PAGE BLANK (USPTO)

L18 ANSWER 9 OF 30 DGENE (C) 2002 THOMSON DERWENT
 AN AAB84360 Protein DGENE
 TI Novel **isolated human kinase** polynucleotide
 that shares structural similarity with animal **kinases** including
 calcium/calmodulin-dependent protein **kinases** and
 serine/threonine protein **kinases**, useful in therapeutics -
 IN Donoho G; Scoville J; Turner C A; Friedrich G; **Zambrowicz B**;
 Abuin A; Sands A T
 PA (LEXI-N) LEXICON GENETICS INC.
 PI WO 2001042435 A2 20010614 32p
 AI WO 2000-US33240 20001207
 PRAI US 1999-169428 19991207
 DT Patent
 LA English
 OS 2001-381667 [40]

=> d his

(FILE 'HOME' ENTERED AT 15:57:10 ON 15 MAY 2002)

INDEX 'ADISALERTS, ADISINSIGHT, ADISNEWS, AGRICOLA, ANABSTR, AQUASCI,
 BIOBUSINESS, BIOCOMMERCE, BIOSIS, BIOTECHABS, BIOTECHDS, BIOTECHNO, CABA,
 CANCERLIT, CAPLUS, CEABA-VTB, CEN, CIN, CONFSCI, CROPB, CROPU, DDFB,
 DDFU, DGENE, DRUGB, DRUGLAUNCH, DRUGMONOG2, ...' ENTERED AT 15:57:19 ON
 15 MAY 2002

SEA HUMAN? AND KINAS?

 475 FILE ADISALERTS
 295 FILE ADISINSIGHT
 54 FILE ADISNEWS
 297 FILE AGRICOLA
 96 FILE ANABSTR
 143 FILE AQUASCI
 407 FILE BIOBUSINESS
 141 FILE BIOCOMMERCE
 76512 FILE BIOSIS
 1664 FILE BIOTECHABS
 1664 FILE BIOTECHDS
 32983 FILE BIOTECHNO
 1347 FILE CABA
 34379 FILE CANCERLIT
 46764 FILE CAPLUS
 112 FILE CEABA-VTB
 33 FILE CEN
 77 FILE CIN
 569 FILE CONFSCI
 27 FILE CROPU
 394 FILE DDFB
 7268 FILE DDFU
 52947 FILE DGENE
 394 FILE DRUGB
 82 FILE DRUGNL
 8737 FILE DRUGU
 123 FILE DRUGUPDATES
 481 FILE EMBAL
 61823 FILE EMBASE
 18573 FILE ESBIODBASE
 28 FILE FROSTI
 29 FILE FSTA
 37193 FILE GENBANK
 15 FILE HEALSAFE
 1105 FILE IFIPAT

THIS PAGE BLANK (USPTO)

5593 FILE JICST-EPLUS
 71 FILE KOSMET
 13635 FILE LIFESCI
 8 FILE MEDICONF
 79649 FILE MEDLINE
 124 FILE NIOSHTIC
 362 FILE NTIS
 18 FILE OCEAN
 20701 FILE PASCAL
 111 FILE PHAR
 1 FILE PHIC
 157 FILE PHIN
 1356 FILE PROMT
 49394 FILE SCISEARCH
 2 FILE SYNTHLINE
 30154 FILE TOXCENTER
 19045 FILE USPATFULL
 55 FILE USPAT2
 1770 FILE WPIDS
 1770 FILE WPINDEX
 34 FILE IPA
 186 FILE NAPRALERT
 1843 FILE NLDB

L1 QUE HUMAN? AND KINAS?

FILE 'MEDLINE, BIOSIS, EMBASE, DGENE, SCISEARCH, CAPLUS, GENBANK,
 CANCERLIT, BIOTECHNO, TOXCENTER, PASCAL, USPATFULL, ESBIODBASE, LIFESCI,
 DRUGU, JICST-EPLUS, NLDB, WPIDS' ENTERED AT 16:00:01 ON 15 MAY 2002

L2 296040 S HUMAN? (S) KINAS?
 L3 73739 S L2 (S) (RECOMBIN? OR ISOLAT? OR CLON?)
 L4 7879 S L3 (S) (SERIN? OR THREON?)
 L5 192 S L3 (S) NEK?
 L6 286 S L3 (S) NY?
 L7 111 DUP REM L5 (81 DUPLICATES REMOVED)
 L8 234 DUP REM L6 (52 DUPLICATES REMOVED)
 E WALKER
 L9 241 S E3
 E TURNER/AU
 L10 1510 S E4
 E TURNER
 L11 64367 S E3
 E WALKER/AU
 L12 0 S E3
 E ZAMBROWICZ/AU
 L13 1088 S E4
 L14 42 S L3 AND (L12 OR L10 OR L13)
 L15 8 S L4 AND (L12 OR L10 OR L13)
 L16 0 S L7 AND (L12 OR L10 OR L13)
 L17 0 S L8 AND (L12 OR L10 OR L13)
 L18 30 DUP REM L14 (12 DUPLICATES REMOVED)

=> log h

COST IN U.S. DOLLARS

SINCE FILE	TOTAL
ENTRY	SESSION
97.00	99.86

FULL ESTIMATED COST

SESSION WILL BE HELD FOR 60 MINUTES
 STN INTERNATIONAL SESSION SUSPENDED AT 16:25:25 ON 15 MAY 2002

THIS PAGE BLANK (USPTO)

**PALM INTRANET**

Day : Wednesday
Date: 5/15/2002
Time: 16:03:10

Inventor Name Search

Enter the **first few letters** of the Inventor's Last Name.
Additionally, enter the **first few letters** of the Inventor's First name.

Last Name**First Name**

(To go back use Back button on your browser toolbar.)

Back to [PALM](#) | [ASSIGNMENT](#) | [OASIS](#) | [Home page](#)

THIS PAGE BLANK (USPTO)

**PALM INTRANET**

Day : Wednesday
Date: 5/15/2002
Time: 16:03:10

Inventor Name Search

Enter the **first few letters** of the Inventor's Last Name.
Additionally, enter the **first few letters** of the Inventor's First name.

Last Name**First Name**

(To go back use Back button on your browser toolbar.)

Back to [PALM](#) | [ASSIGNMENT](#) | [OASIS](#) | [Home page](#)

THIS PAGE BLANK (USPTO)

**PALM INTRANET**

Day : Wednesday
Date: 5/15/2002
Time: 16:03:10

Inventor Name Search

Enter the **first few letters** of the Inventor's Last Name.
Additionally, enter the **first few letters** of the Inventor's First name.

Last Name**First Name**

(To go back use Back button on your browser toolbar.)

Back to [PALM](#) | [ASSIGNMENT](#) | [OASIS](#) | [Home page](#)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 07:55:13 ; Search time 93.14 Seconds
(without alignments)
2254.840 Million cell updates/sec

Title: us-09-783-320-4
Perfect score: 6243
Sequence: 1 MEKTVRLQKIGESFGKAIL.....YAKILHLVWADGAYCEDNDE 1214

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMEL19:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.podent:*
13: sp.yetus:*
14: sp.vertebrate:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6201	99.3	1265	4	Q96PY6
2	3945	63.2	771	4	Q9Y594
3	1953	31.3	375	4	Q9H6P7
4	1387	22.2	320	4	Q96S82
5	836	13.4	744	11	Q95673
6	836	13.4	792	11	Q9R1J1
7	827	13.2	792	11	Q921J2
8	827	13.2	509	11	Q99K72
9	695	11.1	555	5	Q9N9C3
10	622.5	10.0	698	11	Q91ZK4
11	621.5	10.0	606	10	Q9CAU7
12	617	9.9	697	13	Q90XC2
13	614	9.8	291	11	Q9D685
14	609	9.8	943	10	Q94C05
15	596.5	9.6	609	10	Q947T1
16	593	9.5	312	4	Q9Y6S4

17	590.5	9.5	579	5	Q95XQ3	Q95XQ3 caenorhabdi
18	584	9.4	841	5	Q9VC32	Q9VC32 drosophila
19	566.5	9.1	416	10	Q9LR35	Q9LR35 arabidopsis
20	559.5	9.0	561	5	Q76134	Q76134 tetrahymena
21	557	8.9	941	10	Q9LXP3	Q9LXP3 arabidopsis
22	555.5	8.9	443	11	Q91Z18	Q91Z18 mus musculus
23	555	8.9	442	13	Q9W622	Q9W622 xenopus lae
24	548	8.8	384	4	Q960N9	Q960N9 homo sapien
25	546.5	8.8	443	11	Q921N9	Q921N9 mus musculus
26	544.5	8.7	389	13	Q9W653	Q9W653 xenopus lae
27	540.5	8.7	357	5	Q01775	Q01775 caenorhabdi
28	519	8.3	735	5	Q9W41	Q9W41 rattus norv
29	514.5	8.2	723	11	Q91X01	Q91X01 drosophila
30	500	8.0	123	11	Q9K3N8	Q9K3N8 schizosacch
31	494.5	7.9	1257	4	Q9UKP9	Q9UKP9 homo sapien
32	494.5	7.9	1352	4	Q9UKR2	Q9UKR2 homo sapien
33	494.5	7.9	1352	4	Q9UKR3	Q9UKR3 homo sapien
34	493.5	7.9	1202	11	Q9W041	Q9W041 mus musculus
35	490.5	7.8	1305	4	Q9UKD8	Q9UKD8 homo sapien
36	490	7.8	1268	4	Q9UKD6	Q9UKD6 homo sapien
37	490	7.8	1333	4	Q9UKES	Q9UKES homo sapien
38	490	7.8	1235	4	Q9H2G2	Q9H2G2 homo sapien
39	487.5	7.8	1235	4	Q9UKED	Q9UKED homo sapien
40	487.5	7.8	1235	4	Q9UKED	Q9UKED homo sapien
41	486	7.8	1276	4	Q9UKED	Q9UKED homo sapien
42	483.5	7.7	1331	4	Q9UKED	Q9UKED homo sapien
43	482.5	7.7	1204	4	Q00211	Q00211 homo sapien
44	479.5	7.7	306	4	Q9ULX2	Q9ULX2 homo sapien
45	479.5	7.7	313	4	Q96E83	Q96E83 homo sapien

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	1265 AA.
Q96PY6	1	Q96PY6		
AC	Q96PY6	PRELIMINARY	PRT	1265 AA.
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	K1AA1901 PROTEIN (FRAGMENT).			
GN	K1AA1901.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
FN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRAIN;			
RX	MEDLINE=21456161; PubMed=11572484;			
RA	Negase T., Kikuno R., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes. XXI.			
RT	"The complete sequences of 60 new cDNA clones from brain which code for			
RT	large proteins.";			
RL	DNA Res. 8:179-187(2001).			
DR	EMBL; AB067488; BAB67794.1;			
FT	NON_TER			
SEQ	SEQUENCE	1265 AA;	143754 MW;	31DDFD2FEAE39C89 CRC64;

Query Match: 99.3%; Score 6201; DB 4; Length 1265;
Best local Similarity: 96.3%; Pred. No. 6.5e-314;
Matches 1212; Conservative 1; Mismatches 1; Indels 44; Gaps 1;

QY	1	MEKTVRLQKIGESFGKAILVKSTEDGRQYVYKEINISRMSSKREESRRVAVLANKH 60
DB	8	MEKTVRLQKIGESFGKAILVKSTEDGRQYVYKEINISRMSSKREESRRVAVLANKH 67
QY	61	PNIVQYRSPFENSGSLIVMDYCEGDLFRINAGKGVLFQEDQTLDFVQICLAKHVV 120
DB	68	PNIVQYRSPFENSGSLIVMDYCEGDLFRINAGKGVLFQEDQTLDFVQICLAKHVV 127

Wed May 15 14:14:06 2002

us-09-783-320-4.rspt

Page 2

121 DRKILRDIKSONIFLTCKDVTGDLGFIARVNLSTVELARTCTGTPPYLSPECENKPY 180
128 DRKILRDIKSONIFLTCKDVTGDLGFIARVNLSTVELARTCTGTPPYLSPECENKPY 187
181 NMSDITWALGCVLYELCTAKHAFEGASMKNLVLTITSSFPVLSHSTDLRSVQLFK 240
188 NMSDITWALGCVLYELCTAKHAFEGASMKNLVLTITSSFPVLSHSTDLRSVQLFK 247
241 RNPDRPVSNIIEKFIKRIEFLSPOLIAEFCLTKTSKFSQPIPAKRPASGNSI 300
248 RNPDRPVSNIIEKFIKRIEFLSPOLIAEFCLTKTSKFSQPIPAKRPASGNSI 307
301 SVMPAOKITRPAKYGIPLAYKKYDKLHEKKPLQKHQAHOQPEKRVNTGEBRRKISE 360
308 SVMPAOKITRPAKYGIPLAYKKYDKLHEKKPLQKHQAHOQPEKRVNTGEBRRKISE 367
361 EAARKRLEFIEKEKKOKDOITSLMKAEQMKROEKERLERINRAREGMNVLSAGSGE 420
368 EAARKRLEFIEKEKKOKDOITSLMKAEQMKROEKERLERINRAREGMNVLSAGSGE 427
421 VKAPFISGSGTAPSSFSRGQYEHYHAFIDMOQOQRAEDNEAKMKREIYRGLEPE--- 476
428 VKAPFISGSGTAPSSFSRGQYEHYHAFIDMOQOQRAEDNEAKMKREIYRGLEPEGIL 487
477 -----ROKOLAVERAQVEEFLOR 496
488 PGVRPGFYGAAGHHFPPADDIRTKLRKAVSKQANANAKOGLAVERAKOVEEFLOR 547
497 KRAMONKARABGHMYLRLROILRONFNEQOIKAKLGEKKANHSBQESSEADAM 556
548 KRAMONKARABGHMYLRLROILRONFNEQOIKAKLGEKKANHSBQESSEADAM 607
557 RRRKIESLAHANARAVALKEQLEKRRKAEYEREKKEVHEHLVAKGVKSSDVSPILQOHE 616
608 RRRKIESLAHANARAVALKEQLEKRRKAEYEREKKEVHEHLVAKGVKSSDVSPILQOHE 667
617 TGSPSKOQMSVIVTSALKLEVGVDSLTDTRETSEMOKTNNATISSKREILRLNENL 676
668 TGSPSKOQMSVIVTSALKLEVGVDSLTDTRETSEMOKTNNATISSKREILRLNENL 727
677 KAOEDKGMQNLSDTPEIYNVEDAKHEHEKESVSDRRKRWAGQGLVPLDELTLDTSSFS 736
728 KAOEDKGMQNLSDTPEIYNVEDAKHEHEKESVSDRRKRWAGQGLVPLDELTLDTSSFS 787
737 TTRRHVGEYIKLGPNGSPRRAMGKSPDVSYLKILGEABELOLOTELLENTTIRSEISPEG 796
788 TTRRHVGEYIKLGPNGSPRRAMGKSPDVSYLKILGEABELOLOTELLENTTIRSEISPEG 847
797 EKRPILITGEKKVOCISHINPSAIVDSPEVETKSPFESEASPOMSLKEGNLEPDDLET 856
848 EKRPILITGEKKVOCISHINPSAIVDSPEVETKSPFESEASPOMSLKEGNLEPDDLET 907
857 ELIQEPSTGNKDESLPCTITTDWVISEKETKETOSADRITITOEENVEDGVSTVDOLSD 916
908 ELIQEPSTGNKDESLPCTITTDWVISEKETKETOSADRITITOEENVEDGVSTVDOLSD 967
917 IHIEPTGNDOSHCKCDVDSVOPPEPFHKKVYHSEHLNLPVOVSQCSPEESFAFRSHSH 976
968 IHIEPTGNDOSHCKCDVDSVOPPEPFHKKVYHSEHLNLPVOVSQCSPEESFAFRSHSH 1027
977 LPRKNNKNSLILGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMDVPVGVGDVRODNLIED 1036
1028 LPRKNNKNSLILGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMDVPVGVGDVRODNLIED 1087
1037 EIEDENIKGSPSDEIIVEEETDLOLOASMEOLLEBOGGEESSEESVYLSKNSYVER 1096
1088 EIEDENIKGSPSDEIIVEEETDLOLOASMEOLLEBOGGEESSEESVYLSKNSYVER 1147
1097 TANGTVADEDDNPSSSESLNBEHMSDNGDTIASSECDCSVFNHLEELRLHLEQEMKFE 1156
1148 TANGTVADEDDNPSSSESLNBEHMSDNGDTIASSECDCSVFNHLEELRLHLEQEMKFE 1207
1157 KFEFEYERIKAIHEDEDENIETCSKIYONITLGNHQLYAKILHLVMAAGVAYQEDNDE 1214

1208 KFEFEYERIKAIHEDEDENIETCSKIYONITLGNHQLYAKILHLVMAAGVAYQEDNDE 1265
RESULT 2
ID 09Y594 PRELIMINARY; PRT: 771 AA.
AC 09Y594: 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NY-REN-55 ANTIGEN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99438124; Pubmed-10508479;
RA Scanlan M.J., Gordon J.D., Williamson B., Stockert E., Bander N.H.,
Jongeneel V., Gure A.O., Jager D., Jager E., Knuth A., Chen Y.-T.,
Old L.J.;
RT "Antigens recognized by autologous antibody in patients with renal-
cell carcinoma".
RT Int. J. Cancer 83:456-464(1999).
DR EMBL, AF15113, A042879.1; -.
FT NONTER 1
SQ SEQUENCE 771 AA: 87539 MW: 95076131FA32F6CF CRC64:
Query Match 63.28; Score 3945; DB 4; Length 771;
Best Local Similarity 99.68; Pred. No. 5e-197;
Matches 768; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
444 EHYHAFIDMOQOQRAEDNEAKMKREIYRGLEPEROKGOLAVERAQVEEFLORRAMON 503
1 EHYHAFIDMOQOQRAEDNEAKMKREIYRGLEPEROKGOLAVERAQVEEFLORRAMON 60
504 KARAGHMYLRLROILRONFNEQOIKAKLGEKKANHSBQESSEADAMRRKIES 563
61 KARAGHMYLRLROILRONFNEQOIKAKLGEKKANHSBQESSEADAMRRKIES 120
564 LKAHANARAVALKEQLEKRRKAEYEREKKEVHEHLVAKGVKSSDVSPILQOHE 623
121 LKAHANARAVALKEQLEKRRKAEYEREKKEVHEHLVAKGVKSSDVSPILQOHE 180
624 OQMSVIVTSALKLEVGVDSLTDTRETSEMOKTNNATISSKREILRLNENL 683
181 OQMSVIVTSALKLEVGVDSLTDTRETSEMOKTNNATISSKREILRLNENL 743
684 GMQNLSDTPEIYNVEDAKHEHEKESVSDRRKRWAGQGLVPLDELTLDTSSFS 796
241 GMQNLSDTPEIYNVEDAKHEHEKESVSDRRKRWAGQGLVPLDELTLDTSSFS 856
744 GEVILGPNGSPRRAMGKSPDVSYLKILGEABELOLOTELLENTTIRSEISPEG 907
301 GEVILGPNGSPRRAMGKSPDVSYLKILGEABELOLOTELLENTTIRSEISPEG 967
804 TEKKVOCISHINPSAIVDSPEVETKSPFESEASPOMSLKEGNLEPDDLET 1027
361 TEKKVOCISHINPSAIVDSPEVETKSPFESEASPOMSLKEGNLEPDDLET 1087
864 GNNKRESLPCTITTDWVISEKETKETOSADRITITOEENVEDGVSTVDOLSD 1147
421 GNNKRESLPCTITTDWVISEKETKETOSADRITITOEENVEDGVSTVDOLSD 1207
924 NDSQSKCDVDSVOPPEPFHKKVYHSEHLNLPVOVSQCSPEESFAFRSHSH 1265
481 NDSQSKCDVDSVOPPEPFHKKVYHSEHLNLPVOVSQCSPEESFAFRSHSH 1325
984 KNSLILGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMDVPVGVGDVRODNLIED 1385
541 KNSLILGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMDVPVGVGDVRODNLIED 1445

```

OY 1044 KEGPDSSEDIIVFEETDTLOELQASMEQLLRQEPGESESEESVLSKNSDVEPTANGTDV 1103
      |||||||
DB 601 KEGPDSSEDIIVFEETDTLOELQASMEQLLRQEPGESESEESVLSKNSDVEPTANGTDV 660
OY 1104 ADEDDNPSSSSALNEEMHSDNSNGEFASECDSVFNHLEELRLHLEQEMGEFFFEVYE 1163
      |||||||
DB 661 ADEDDNPSSSSALNEEMHSDNSNGEFASECDSVFNHLEELRLHLEQEMGEFFFEVYE 720
OY 1164 KIKAHEDBENETICSTYONILNENHOLYAKITILVMDAGAYOEENDE 1214
      |||||||
DB 721 KIKAHEDBENETICSTYONILNENHOLYAKITILVMDAGAYOEENDE 771

RESULT 3
OY 09H6P7 PRELIMINARY; PRT: 375 AA.
AC 09H6P7;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE CDNA: FLJ22005 FIS, CLONE HEP06902.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID:9606;
RN
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK025658; BAB15207.1; -
SO SEQUENCE 375 AA; 42401 MW; 97A51B0140F9DD09 CRC64;

Query Match 31.3%; Score 1953; DB 4; Length 375;
Best local similarity 99.7%; Pred. No. 5,6e-94;
Matches 374; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 840 MSIKLEGNLEFPDLEETLOEPSTGNKDESLPCTITDVISSEKKEKQOSADRTTIOE 899
      |||||||
DB 1 MSIKLEGNLEFPDLEETLOEPSTGNKDESLPCTITDVISSEKKEKQOSADRTTIOE 60
OY 900 NEYSEGVASTYDOLSDHIEGCTNDSDHSCDYKSVOPPEPHKVVHSEHLNLPVOY 959
      |||||||
DB 61 NEYSEGVASTYDOLSDHIEGCTNDSDHSCDYKSVOPPEPHKVVHSEHLNLPVOY 120
OY 960 SVOCSPESFAFRSHSLPPKNNKNSLLIGSLGLDANNPKMLRTCSLPDLSKLEFRL 1019
      |||||||
DB 121 SVOCSPESFAFRSHSLPPKNNKNSLLIGSLGLDANNPKMLRTCSLPDLSKLEFRL 180
OY 1020 MDVPTGVADYRQNDLEIDELIKDENIKESPSDEDIIVFEETDTLOELQASMEQLLRQEPGE 1079
      |||||||
DB 181 MDVPTGVADYRQNDLEIDELIKDENIKESPSDEDIIVFEETDTLOELQASMEQLLRQEPGE 240
OY 1080 EYSEEESEVLSKNSDVEPTANGTDVADEDDNPSSSSALNEEMHSDNSNGEFASECDSVF 1139
      |||||||
DB 241 EYSEEESEVLSKNSDVEPTANGTDVADEDDNPSSSSALNEEMHSDNSNGEFASECDSVF 300
OY 1140 NHLLEELRLHLEQEMGEFFFEVYEYKIKAHEDBENETICSTYONILNENHOLYAKITL 1199
      |||||||
DB 301 NHLLEELRLHLEQEMGEFFFEVYEYKIKAHEDBENETICSTYONILNENHOLYAKITL 360
OY 1200 HLVMADGAYOEENDE 1214
      |||||||
DB 361 HLVMADGAYOEENDE 375

RESULT 4
OY 096S52 PRELIMINARY; PRT: 320 AA.
ID 096S52

```

```

AC 096S52;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CDNA: FLJ14674 FIS, CLONE NT2RP2003912, HIGHLY SIMILAR TO
DE SERINE/THREONINE-PROTEIN KINASE NEK1 (EC 2.7.1.-).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID:9606;
RN
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsumura H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuh Y., Sasaki N.;
RT "NEO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK027580; BAB5209.1; -
SO SEQUENCE 320 AA; 36509 MW; F64FE9BB69C0F730 CRC64;

Query Match 22.2%; Score 1387; DB 4; Length 320;
Best local similarity 100.0%; Pred. No. 1.1e-64;
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 208 MKNLVLIKIIISGFPVSYHYSDIRSLVQLFKRNPDRPSVNSILEKGFIAKRIEKL 267
      |||||||
DB 1 MKNLVLIKIIISGFPVSYHYSDIRSLVQLFKRNPDRPSVNSILEKGFIAKRIEKL 60
OY 268 POLIAFEELKTFKFSGQPIPAKRPASGONSISVMPAQKITPAKYGIPLAYKKYGD 327
      |||||||
DB 61 POLIAFEELKTFKFSGQPIPAKRPASGONSISVMPAQKITPAKYGIPLAYKKYGD 120
OY 328 KLEHKKPLOKHQAOHPKRVNGEERKISEAAKRRLEFEKKEKOKDQIISLMA 387
      |||||||
DB 121 KLEHKKPLOKHQAOHPKRVNGEERKISEAAKRRLEFEKKEKOKDQIISLMA 180
OY 388 EOKKROEKRLERINARPOGHRNLSAGSGEYKAPFLSGGTIAPSSSSSGOYEYH 447
      |||||||
DB 181 EOKKROEKRLERINARPOGHRNLSAGSGEYKAPFLSGGTIAPSSSSSGOYEYH 240
OY 448 AIFDQMOQOQRAEDNEAKKRELYGGLPER 477
      |||||||
DB 241 AIFDQMOQOQRAEDNEAKKRELYGGLPER 270

RESULT 5
ID 035673 PRELIMINARY; PRT: 744 AA.
AC 035673;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-DEC-2001 (Tremblrel. 05, Last sequence update)
DE MSTRK25 KINASE-LIKE PROTEIN.
DE NEK4 OR MSTRK25.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN
RP SEQUENCE FROM N.A.
RA STRAIN-BALB/C;
RA MEDLINE:20001940; PubMed:10529384;
RA Hayashi K., Igarashi H., Ogawa M., Sakaguchi N.;
RT "Active and substrate specificity of the murine STRK2
RT regulator protein NIMA of Aspergillus nidulans.";
RL Biochem. Biophys. Res. Commun. 264:449-456(1999).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: Y09234; CAI0436.1; -
DR HSSP: 063450; 1A06.
DR MGI:1344404; Nek4.

```

Wed May 15 14:14:06 2002

us-09-783-320-4.rspt

Page 4

DR InterPro: IPR0000719; Euk_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR Pfam: PF00069; kinase_1.
DR SMART: SM00220; S_TKc_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50108; PROTEIN_KINASE_ST_1.
DR ATP_binding; Kinase; Serine/threonine/protein kinase; Transferase
KW SEQUENCE 744 Aa; 83542; 82E1AAFE39CC72A CRC64;

Query Match	Similarity	30.1%	Score 836	DB 11	Length 744
Best Local	Similarity	30.1%	Pred. No. 1,3e-35		
Matches	226	Conservative	128	Mismatches	261
				Indels	160
				Gaps	24
QY	4	YVFLKQIGESSECKALIVSTEDGROYVYIKENISRMSKRESREKREAVLANMKHPT	63		
DB	6	YCMRVVAGKRSIGEVTVLHNRDQKQYVYIKNLNNSSRRRAEDQALQSLKPR	65		
QY	64	VOYRESEF-ENGSLYVMDYCEGGLDEFRINQGVLFQEDQIIDMFVQICLAKHYDR	122		
DB	66	VYTKESMEGGDGLTYVMGFCGGGLYKRLKEQKQQLLPESQVVEWVQIYAMALQYIHEK	125		
QY	123	KILHARDIKSONIEFLTKDGTVOQDEGIARVINSTVETLARTICIGTPYVLSPEIDENKPYN	182		
DB	126	HIHLRPLKQONVLEFTRTNIIKVGDDIGIARVLENNHDMASLTICGPIYVMSPELSPKNRYN	185		
QY	183	KSDIMALGCVLELCLKHAFAGSMKMLVYKIIISGSPRPVSLHYSVDLSLSQLEKRN	242		
DB	186	KSDVMAIGCCVYEMALIKHAFNAKMNSLVYRIIEGKLPMPKRYSTELMELLRTMLSR	245		
QY	243	PRDRPSVNIILEKGFATKRIEKFSLPOLIAEFCLCTEFSKFG-----SADIPAKRPAS	295		
DB	246	PERPSPVSIILRPYIKIHISLFL-----EATKAKT-SKNVKNKDSRAKPAVAAYSRK	298		
QY	296	GONSISVMPAKOTTRKAAKYGTPLVKKYGGKCIHEKKPLQ-----KHKOAHQPREK	348		
DB	299	EESNTVDYIHQ-----PRSEGSAL-IHWGDEKCLSQEKPVDIGLRSPLSIEGHTQKDM	353		
QY	349	VNNGEERKRISELAARKRLFEIEKK-----QKDIQIISLAKAEQK-----ROEK	395		
DB	354	NNGEESCATI-----RINIDILPERKRSANAGVYSDSQOHADADVEDVDSQSIQEK	408		
QY	396	ERLEIRINARQEGIMRVL-----SAGGSG-----EVKAPFLGSGGTAPSSFSRGO--YEHY	446		
DB	409	ERLDGNTKSSDQ--GULLPRRSSDGDGSELYKPLVPSKKDKKPPQFELPRLLPSY	467		
QY	447	HAI--FDQMOQQAEDNE-----AKMKREIYGRGILPEROKQLAVER	486		
DB	468	PGIGKVDIATQONDNOGPGVAGCVNSRTSSTASAK-----DPLSARER-----RR	516		
QY	487	AKVVEEFLQKRREAMONKARAGHNYUATLRIQLRIQNFN-----ROQIKAK	534		
DB	517	LKQSOEEMLRPGSPAVOVRTSAVERLKPQEDOPRIQAQFSSDSCSITQMNHTLPREREKL	576		
QY	535	LRG-EKKEAHSNGCGSEADMRKKITESTKANHANAAYLAKQOLEKKREAYE-----	588		
DB	577	MHGLSEDELSSSTSDKSDGDSRSG-----KSHNEMKDLVOLMTQTLRLKESGEDL	631		
QY	589	-----REKVVMEHNLVAKGVSSDPYSLGHEHSGSPSKOBSYVTSYLSALKEY	639		
DB	632	QVNLPGSEFRLHKKYDVLVLYGKAAVEYERPCSTELPGITIPSEKIRIYEVURA-----	667		
QY	640	GVDSDLTDRETISEEMQKTNNAISSKREILRLRLNENLKAODEKGMONLSDTEINVED	699		
DB	688	-----DVIQGL-----GIQLLEQVDFDLGED	709		
QY	700	AKENE 704			
DB	710	ELERE 714			

RESULT 6

```

09RLJ1
ID 09RLJ1          PRELIMINARY;          PRT;          792 AA.
AC 09RLJ1;
DT 01-MAY-2000 (TReMBLrel_13, Created)
DT 01-MAY-2000 (TReMBLrel_13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel_19, Last annotation update)
DE SERINE/THREONINE-PROTEIN KINASE NEK4.
GN NEK4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteleia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=99321807; Pubmed=10393247;
RA Chen A., Yanai A., Arima E., Kilfin G., Motro B.;
RT "NIMA-related kinases: Isolation and characterization of murine nek3
RT and nek4 cDNAs, and chromosomal localization of nek1, nek2 and nek3."
RL Gene 234:127-137(1999).
CC EML: SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC EML: AF099067; AAC16287.1; .
DR HSP: O63450.1A06.
DR MGD: MGI:1344404; Nek4.
DR InterPro: IPR000719; Puk.pkinase.
DR InterPro: IPR002390; Ser_thr.pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; Pkinase.1.
DR PRINTS: PRO0109; TPKINASE.
DR SMART: SM00220; S_TKc_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST_1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 792 AA; 89093 MW; 9B8F37117377CC6 CRC64.

```

Query Match	13.4%	Score 836	DB 11	Length 792
Best Local Similarity	29.0%	Pred. No. 1.5e-35		
Matches	239	Conservative	129	Mismatches 269; Indels 186; Gaps 22

QY	4	YVRLOKIGESPKALIVSTEGRGVYVKEINISRMSSKRESESRREVALVANKHPNI	63
DB	6	KYKRVAVGRSGYEVLVLYKHRDRGKQYVTKIKLNLNASSRRRAAEQADQLSOLKHPNI	65
QY	64	VQYEESE-ENGSLIYMDYCEGGDLFKYINAKOVLQEOHLLMDVYOLCAKLVHVR	122
DB	66	VTYKESDEGGDLIYMGFCBEGDLYKRLKQKQKQDLPEQVYEWVYQIYAMLYQLHKK	125
QY	123	KILHRDIKSONIFLTKDGYVLQDGSIAVNLVSLYELARCTICIPYTLSEIENCPYNN	182
DB	126	HLHRDLKTQWVEFLTRNIYKIGDDLIAYLVNHDGMASLTILGRPLYMSELFSNKPYN	185
QY	183	KSDIWAIGCVLYELCTIKNAFEAGSKKNLYLKSISPPVSLHYSTDLASLYSQLFKRN	242
DB	186	KSDWALGACCCYEATLTKNAFNKDKMNSLVYIITLGGKLPAPKVVYSTLLEMLIRTLMSR	245
QY	243	PRDRPSVNLIEKGFIAKRIEKLFSQDLAEFECLKTFSKFG-----SQTIPAKRPAS	295
DB	246	PEERPSVSLIRLOPYLKHNLSTFL-----EYTKAKT-SKNNVKNCDSRAKPAALVASKR	298
QY	296	GONSISVMPAKITTKPAKAGYIPLAVKKGDKKLEKKRPLO-----KKHQAHQTPKKR	348
DB	299	EESNTVDYIHQ-----PRSSGSAFL-HVMGEDCLSQEQRVPDVLGFLRSPASLEHGTGKOM	353
QY	349	VNTEEYERRKISELAKARRRLEFLEKEKK-----QKQDLISLMKAQOMK-----ROEK	395
DB	354	NNTSESCATIS-----RINIDILPERRRBSANAGVVDSPQHNDADDEVDSDCSISQEK	408
QY	366	ERLRINIRAREQGRWVL-----SAGSGG-----EVKAPFEGSGGTIAPSPFSFSROUYEHYA	448
DB	409	ERLQGNKTSKSDOP-GNLLPRSSDGGDGGSELVPLTPSPNKKDQKPPQDOVVTGIIENQDS	467
QY	449	IFPDQMO-----QQRADDEAKWKRREITYGRQLPE-----	476

```

Db 468 IHPSSQPHSSMSEPSLSRQRRKRRKQTASHSTKSOPDELPLPLPSYIGKVDIATQ 527
Oy 477 ---ROKQOLA-----VERKAYEEFLORREAROMKARREG 509
Db 528 QNDGOGGPPVAGVNSSRTSYASAKDRPLSARERRRLKQOSEEMLPSPGPAVOPTPSAVE 587
Oy 510 HMYVLARLROIRLONENE-----ROQIKAKLRG-EKKANHSQEGSGSEADM 556
Db 588 PLKPEOEOPTPAQFSSDCSTOMNHTLPREKRLMGLSLEDLSSTSTOKSDGS 647
Oy 557 RRRKIESLKAHANAAVALEKQERKREAYE-----REKVMEEHLVAKG 602
Db 648 REG-----KSHTNEMKDLVOLMTOTLRLKAKSCEDLOVLPNGSEFRLHRYKROTLYLHG 702
Oy 603 VKSDVSPPLGQHETGSGSPSKQOMRSVTSALKVEGVDSLTIDRTSEEMKTNNAI 662
Db 703 KVAEEVEPHCTELPTGITPGSEKIRIRIVEVLA-----735
Oy 663 SSKREILRLNENLKAODEKGMONLSDTFEINVEDAKENE 704
Db 736 ---DVIOGL-----GIQLLEOVFDLGEDELERE 762

RESULT 7
ID 0921J2 PRELIMINARY: PRT: 792 AA.
AC 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DR 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SERINE/THREONINE KINASE PROTEIN MSTK2L LONG-FORM.
GN NEK4 OR MSTK2L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RC STRAIN-BAB/C; TISSUE-SPLEEN;
RX MEDLINE=20001940; PubMed=10529384;
RA Hayashi K., Igarashi H., Ogawa M., Sakaueuchi N.;
RT "Activity and substrate specificity of the murine STK2
RT serine/threonine kinase that is structurally related to the mitotic
RT regulator protein NIMA of Aspergillus nidulans."
RL Biochem. Biophys. Res. Commun. 264:449-456(1999).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AJ223071; CAAL1072.1;
DR HSSP: Q63450; IAO6.
DR MGD: MGI:1344404; Nek4.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002290; Ser_thr_Pkinase.
DR InterPro: IPR001245; Tyr_Pkinase.
DR Pfam: PF00069; TYRKINASE.
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 792 AA: 88994 MW; CE9187311C807A1D CRC64;

Query Match 13.4%; Score 836; DB 11; Length 792;
Best Local Similarity 29.0%; Pred. NO. 1.5e-35; Indels 186; Gaps 22;
Matches 238; Conservative 129; Mismatches 269;

Oy 4 YVRLOKIGESFGKALIVSTEDGROYVKEINISRMSSKREESREAVLANMKHPNI 63
Db 6 KCYRVVGRSGYGVTLVKHRRGKQYVIKKILNRNASSRRRAAEDAOQLSOLKHPNI 65
Oy 64 VQYSESEF-ENGSLIYVNDCEGDDFFKRNKQGVLEQEDQIDMFVQICLAKVHDR 122

```

```

Db 66 VTYKESWEGSDGLLYVMGFCGEGDLYRKLKQKQGLLPESQVVEFMVQIAMAQYLERK 125
Oy 123 KILHROKSONIFLTFEDGTVQVQDGFIARVINSVYELARICIGTPYVLSPEICENKPYNN 182
Db 126 HILHRDLKTONVLTITNIIKVGDIARVLENBDMASITIGIPYMSSELSSNKPYN 185
Oy 183 KSDIAMAIGCYLYELCTKHAFFENGSMKNLVLTIGSGSPVSLYSDLSVLSQLEKRN 242
Db 186 KSDVWALGCVYEBATLKAHFNKDMNSLVYRIEGKLPMPKYSTLELARTMISRR 245
Oy 243 PRDRPSVNSITLKEFTAKRTEKFLSPQILAEFECLKTSKFG-----SQTIPAKRPAS 295
Db 246 PEERSVNSILKOPYIKHHISLFL-----EATKAKT-SKNVKNCSDRAPVAAYVSRK 298
Oy 296 GONSTISVMPAKITKPAKKGITPAKKGIDKILHEKKPLQ-----KHOAHQTPPKR 348
Db 299 EESNNDVVIHYO-----PSSSEKAL-HYMGEDKLSQEKVVDIGPLRSPASLEGHTGKDM 353
Oy 349 VNTGERRKISEEAKRRLEFEKKEK-----OKQOITSIMKAKDM-----ROEK 395
Db 354 NNTGSCATIS-----RINDILPAENRDSANAGVQESOPQHVDAADRVDSOCISQEK 408
Oy 396 ERLERINRAEGQWNL-----SAGSG---EYKAPFLSGTIAPSFSRSGOYEHYHA 448
Db 409 ERLGNTKSSDOP-GNLLPRSSDGGGSGSELVKPLTPSNKDKRPDQDQVGTIENQDS 467
Oy 449 IFDQWQ-----QPARDNEAKRKRELYGGLPE-----476
Db 468 IHPSSQPHSSMSEPSLSRQRRKRRKQTASHSTKSOPDELPLPLPSYIGKVDIATQ 527
Oy 477 ---ROKQOLA-----VERKAYEEFLORREAROMKARREG 509
Db 528 QNDGOGGPPVAGVNSSRTSYASAKDRPLSARERRRLKQOSEEMLPSPGPAVOPTPSAVE 587
Oy 510 HMYVLARLROIRLONENE-----ROQIKAKLRG-EKKANHSQEGSGSEADM 556
Db 588 PLKPEOEOPTPAQFSSDCSTOMNHTLPREKRLMGLSLEDLSSTSTOKSDGS 647
Oy 557 RRRKIESLKAHANAAVALEKQERKREAYE-----REKVMEEHLVAKG 602
Db 648 REG-----KSHTNEMKDLVOLMTOTLRLKAKSCEDLOVLPNGSEFRLHRYKROTLYLHG 702
Oy 603 VKSDVSPPLGQHETGSGSPSKQOMRSVTSALKVEGVDSLTIDRTSEEMKTNNAI 662
Db 703 KVAEEVEPHCTELPTGITPGSEKIRIRIVEVLA-----735
Oy 663 SSKREILRLNENLKAODEKGMONLSDTFEINVEDAKENE 704
Db 736 ---DVIOGL-----GIQLLEOVFDLGEDELERE 762

RESULT 8
ID 099K72 PRELIMINARY: PRT: 509 AA.
AC 099K72:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DR 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SIMILAR TO NIMA (NEVER IN MITOSIS GENE A)-RELATED EXPRESSED KINASE
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RC SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC EMBL: BC005411; AAH05411.1;
DR HSSP: Q63450; IAO6.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002290; Ser_thr_Pkinase.

```

DR InterPro: IPR001245; Tyr-kinase.
 DR Pfam: PF00069; pkinase.1.
 DR PRINTS: PRO0109; TYRKINASE.
 DR SMART: SM00220; S_TKC.1.
 DR SMART: SM00219; TYRK.1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
 DR PROSITE: PS50011; PROTEIN KINASE DOM.1.
 DR PROSITE: PS00108; PROTEIN KINASE ST.1.
 DR ATP-binding: kinase; Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 509 AA; 57109 MW; 335850588686A1 CRC64;

Query Match 13.28; Score 827; DB 11; Length 509;
 Best Local Similarity 37.78; Pred. No. 2.4e-35;
 Matches 204; Conservative 76; Mismatches 175; Indels 86; Gaps 16;

QY 1 MEKYVRLQKIGESFSGKALIVKSTEDGRQYVKEINISMSKRESESRREAVLANMKH 60
 DB 1 MONTYVRLVYGGSGFGRLALVLOESSNOTFAMKEI---RLKSDTOTSREAVLANMKH 57
 QY 61 PNIVQYRESFEENGSLYIVMDYCEGDLFKRINAQGVLFQEDQILDMFVOICLAKHYH 120
 DB 58 PNIVAFKESFEAGLYIVMEYCDGDLQRIKQKGLPREDTILMWFQICLGYNH 117
 QY 121 DRKILHRDKSONIFLTQDGTVOAGDGIARVNLSTVELARTCIGPPYLSPEICENKPY 180
 DB 118 KRRVLRDRKSNVFLTHNGVKVGLGDEGSARLLSSPMAFCTYVGPPIYPPIMENLRY 177
 QY 181 NKSQDIMALGVLYELCTLKAFAEGSMKNLYLKIISGFPPVSLHYSYDLBSYLSOLK 240
 DB 178 NKSQDIMALGVLYELCTLKAFAEGSMKNLYLKIISGFPPVSLHYSYDLBSYLSOLK 237
 QY 241 RNPDRPSPVNSILEKGFIAKRIEFLSPQLAE-----EPCLTFKSGSGPIPAKRP 293
 DB 238 RNSHRPSTYTLICRGSLAPVLKCLPQIIRREYGEIDLEIKISTPKNNKKDSNRVR 297
 QY 294 ASGQ-NSISVMPAQK-----TYKPA-----AKYGIPLAKKYYDKKLEKKP 334
 DB 298 ALGEANASASMOEERGRKCSHTELESTGTTPAGNALRAARGNGNQDEIG---RHSP 353
 QY 335 LQHKQAQGT--PEKRVN-----TGEERRKIS---EAAARKRLFEIEKKK 376
 DB 354 ASHPHPRWRHGRPSSNVALEKASILLSSFAEDDRGSGYIKYEENARQWV-----RE 407
 QY 377 OKDOIISLKAQOMKROKKEKLERLINRAREQGNRVLSAGSGGVKAPFGSGGTAPSS 436
 DB 408 PEPALLSMKLDLDSQAFQT-----YTIYRPGAGELGKPL--SEDTASDSV 452
 QY 437 FSSRGY---EHYHAFIDQMOQOQRAEDNE-AKKKREI-----YGRGLPRROKGLAVR 486
 DB 453 DDDLDVMDPRFRPRPRDEEDTDEEDENENPDLSELKKNVGYGDS---PGQLLGEC 508
 QY 487 A 487
 DB 509 A 509

RESULT 9
 ID Q9N9C3 PRELIMINARY; PRT; 555 AA.
 AC Q9N9C3;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DE NEK-RELATED SERINE/THREONINE-PROTEIN KINASE NEK1.
 GN P1408.02.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RA Botte G., Pohl T., Ivens A.C., Quail M., Rajandream M.A.,

RA Barrell B.G.;
 RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
 RX [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RX MEDLINE=98146435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome."
 RL Genome Res. 8:135-145(1998).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AL358652; CAB94013.1; -.
 DR HSSP: Q00534; 1818.
 DR InterPro: IPR000719; Euk-kinase.
 DR InterPro: IPR002290; Ser-thr-kinase.
 DR InterPro: IPR001245; Tyr-kinase.
 DR Pfam: PF00069; pkinase.1.
 DR PRINTS: PRO0109; TYRKINASE.
 DR SMART: SM00220; S_TKC.1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
 DR PROSITE: PS50011; PROTEIN KINASE DOM.1.
 DR PROSITE: PS00108; PROTEIN KINASE ST.1.
 KW ATP-binding: kinase; Serine/threonine-protein kinase; Transferase.
 SO SEQUENCE 555 AA; 62230 MW; 98054e57d4b39849 CRC64;

Query Match 11.18; Score 695; DB 5; Length 555;
 Best Local Similarity 40.28; Pred. No. 1.9e-28;
 Matches 131; Conservative 85; Mismatches 98; Indels 12; Gaps 3;

QY 1 MEKYVRLQKIGESFSGKALIVKSTEDGRQYVKEINISMSKRESESRREAVLANMKH 60
 DB 1 MDKTYIKRKNIGKNGMGTCTLANNEGKTYIKQYDLTRMSKKRQSLNDAVLYSLRH 60
 QY 61 PNIVQYRESF--EENGSLYIVMDYCEGDLFKRINAQGVLFQEDQILDMFVOICLAKH 118
 DB 61 PNIIYVDSVLAKKSDNLCTVMEVAGSDVCTRLKKNHGVNPPROYVMDLIDLVSLD 120
 QY 119 VHRKILHRDKSONIFLTQDGTVOAGDGIARVNLSTVELARTCIGTPYLSPEICENK 178
 DB 121 VHQRTILHRVKTQNTFLTHENMLIKLDFGIARTLANYYDOAQTFVGTPLYLSEELLE 180
 QY 179 PYNKSDIMALGVLYELCTLKAFAEGSMKNLYLKIISGFPPVSLHYSYDLBSYLSOL 238
 DB 181 PYDRSPVVALGVLYELCTLKAFAEGSMKNLYLKIISGFPPVSLHYSYDLBSYLSOL 240
 QY 239 FKNPDRPSPVNSILEKGFIAKRIEFLSPQLAEFCLTFKSGSGPIPAKRPASGVN 298
 DB 241 LVNDPAGRIKLDLIDLPYVRERIRQWLKEPPVYQNHVYRSLCHNHLF-----DFOD 293
 QY 299 SISVMPAQKTKRAKYGPIPLAKKY 324
 DB 294 EATMPS--TRSAARAAAAMAAHEW 316

RESULT 10
 ID Q912R4 PRELIMINARY; PRT; 698 AA.
 AC Q912R4;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DE NIMA-RELATED KINASE 8.
 GN NEK8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C37BL/6J;
 RA Liu S., Lu W., Obara-Ishihara T., Drummond I., Beier D.R.;
 RT "A defect in a novel Nek-family kinase causes cystic disease in the

RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Murai A.,
RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada H., Yasuda M., Tabata S.:
RT "Sequence and analysis of chromosome 3 of the plant *Arabidopsis*
RT *thaliana*." and
RL Nature 408:820-823 (2000).
RL EMBL, AC009465.1, AACG51423.1, -.
DR HSSP, Q63450.1, IAO6.
DR InterPro: IPR0000719: Euk_pkinase.
DR InterPro: IPR002885: PPR.
DR InterPro: IPR002290: Ser_thr_pkinase.
DR InterPro: IPR001245: Tyr_pkinase.
DR InterPro: P000065: Pkinase.1.
DR Pfam, P000220: S_TKC.1.
DR SMART, SM00219: TYRKC.1.
DR SMART, PS00107: PROTEIN KINASE_ATP; UNKNOWN.1.
DR PROSITE, PS50011: PROTEIN KINASE_DOM.1.
DR PROSITE, PS50011: PROTEIN KINASE_TRANSFER.
KW ATP-binding; kinase; Transferase.
KW SF000006 606 AA; 68034 MW; 18733726C62E3B9C CRC64;
SQ

Wed May 15 14:14:06 2002

us-09-783-320-4.rsp

Page 10

```
QY 297 QNSISVMPAKITRPAKXGILPAAKXGDKKLEHKKPLOKHOAHQTPKERNVTGEER- 355
Db 283 DNDY-----KKTIFA-----ISESTPVKKNR-----EKRMSGNDRT 314
QY 356 -----RKISEBA--ARKRRLEFIEKE 374
Db 315 LNPVSDDHDTFESNRRYPKTPSRVSELSVSPDRGSTVTKKITSKALLVNRNPQIVPKL 374
QY 375 KKOKDOIISLMKAEQKKROEKER-----LERINRAREQWNRNVLASAGSGEVKAPF 425
Db 375 TTPPARQVLDLRNSDMASFTLVKRSVSTTRRASLPLTNKAAVQELPRRPSLSFLDCIKSPD 434
QY 426 LGSGGTIAPS-----SESSRGQVEHYHAFDQMOQORADNEAKWKREITVGRGLPERQKG 480
Db 435 VSVN---APRIDKMLEFFPLASTYEDPFHPI-----RTSSNSA-----QSSSGSPOA 477
QY 481 QLAVERAKOYEFEFLORRREAMQNK---ARAGSHVYTLARLOIRLONFNERQOIKAKLRG 537
Db 478 EYSVMKDKCTIOTPSKFDPMSSNDAMOGEGPMVHV-----DREDIT----- 520
QY 538 EKKRANHSEGOEGSEADMRKK--TESLKAHANARAAVLKEQLEKRRKREAYE----- 588
Db 521 DSSDONATAGASSRTSSDTRRRRRFNNSSYKQRAEALGELLEFSARLLQEDRFDELGVLLK 580
QY 589 -----REKKVMEEHVLAKGVKSS 606
Db 581 PGPEKVSPPRETAIW-----LTKSIKEN 603
```

Search completed: May 15, 2002, 08:01:42
Job time: 389 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 07:51:30 ; Search time 56.63 seconds
(without alignments)
2059.906 Million cell updates/sec

Title: US-09-783-320-4
Perfect score: 6243
Sequence: 1 MEKYVRLQKIGEGSFGKAIL.....YAKILHLVMDAGVQEDNDE 1214

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-71:
1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3385	54.2	774	2	S25284	protein kinase nek
2	847.5	13.6	841	1	I78885	serine/threonine-s
3	836	13.4	792	2	JC7122	protein kinase (EC
4	565.5	9.1	431	2	T11854	protein kinase-lik
5	557	8.9	941	2	T49136	NIMA-like protein
6	550	8.8	445	2	G01452	hypothetical prote
7	540.5	8.7	357	2	T29771	probable protein k
8	534.5	8.5	699	2	A43734	probable G2-specif
9	514.5	8.2	722	2	T37970	G1N4 protein - yea
10	502.5	8.0	1233	2	S59359	serine/threonine p
11	500	8.0	1233	2	T14157	NIMA-like protein
12	499.5	8.0	779	2	A57177	hypothetical prote
13	475	7.6	200	2	B96587	protein kinase - s
14	467	7.5	1094	2	S49313	protein kinase SK2
15	464.5	7.4	1206	2	T34021	hypothetical prote
16	462	7.4	1211	2	T27522	hypothetical prote
17	455.5	7.3	1231	2	T18532	serine/threonine p
18	455	7.3	1354	2	S74244	serine/threonine-s
19	452	7.2	1518	2	S37928	probable purine nu
20	451.5	7.2	1233	2	T30989	serine/threonine p
21	451	7.2	1558	2	T29253	hypothetical prote
22	445	7.1	1354	2	S69211	serine/threonine-s
23	437.5	7.0	915	2	S74383	probable protein k
24	432	6.9	435	2	S23580	serine/threonine-s
25	427	6.8	1388	2	S74245	serine/threonine-s
26	425	6.8	1388	2	S70633	serine/threonine-s
27	422	6.8	1732	2	T14039	protein kinase (EC
28	421.5	6.8	1895	2	T15881	hypothetical prote
29	418.5	6.7	294	2	T21075	hypothetical prote

30	407.5	6.5	925	2	A55748	protein kinase (EC
31	405	6.5	1001	2	T17365	serine/threonine p
32	402	6.4	740	2	T24340	hypothetical prote
33	402	6.4	797	2	T23927	hypothetical prote
34	400	6.4	1702	2	T14050	protein kinase (EC
35	399	6.4	1114	2	T38224	protein-serine/thr
36	398.5	6.4	1246	2	G89287	protein H39E23.1
37	393	6.3	1080	2	S48944	hypothetical prote
38	393	6.3	1173	2	T25539	hypothetical prote
39	390.5	6.3	1192	2	T18611	probable serine/th
40	388	6.2	1062	2	S46367	hypothetical prote
41	387.5	6.2	1228	2	T18697	hypothetical prote
42	383.5	6.1	465	2	B57748	protein kinase (EC
43	383.5	6.1	705	2	A48144	protein kinase CDC
44	382.5	6.1	740	2	T38536	ribosomal protein
45	382	6.1	713	2	S27966	probable serine/th

ALIGNMENTS

RESULT 1
S25284
protein kinase nek1 (EC 2.7.1.-) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Sep-1993
C:Accession: S25284
R:Netrin K.; Mizzen, L.; Mottro, B.; Ben-David, Y.; Bernstein, A.; Pawson, T.
EMBL J. 11, 3521-3531, 1992
A:Title: A mammalian dual specificity protein kinase, Nek1, is related to the NIMA ce
A:Reference number: S25284; M01D:93010942
A:Accession: S25284
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-774 <LEP>
A:Cross-references: GB:S45828; NID:g256854; PIDN:AAB23529.1; PID:g256855
A:Gene: nek1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: ATP; phosphotransferase
F:2-258/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif

Query Match	Best Local Similarity	Score	DB 2;	Length	774;
Matches	665;	Conservative	45;	Mismatches	4;
Gaps					3;
QY	1	MEKYVRLQKIGEGSFGKAILVSTEDGROYVYKEINISRMASRKRESREAVLANMKH	60		
DB	1	MEKYVRLQKIGEGSFGKAVLVSTEDGRHYVKEINISRMASRKRESREAVLANMKH	60		
QY	61	PNIVQYSEFENGSLYIVMDYCEGGDLFRINAQGVLFQEDQILDMFYQICALKHYH	120		
DB	61	PNIVQYSEFENGSLYIVMDYCEGGDLFRINAQGVLFQEDQILDMFYQICALKHYH	120		
QY	121	DRKILHRIKSONFFLRKDGTVOLGDFGIARVNSTVETLARTIGTPYVYSPEICENKY	180		
DB	121	DRKILHRIKSONFFLRKDGTVOLGDFGIARVNSTVETLARTIGTPYVYSPEICENKY	180		
QY	181	NNKSGIWMIGCVLVECTLKHAFEGASMKNLVLIISGSPVSLHVSYDLSLSOLF	240		
DB	181	NNKSGIWMIGCVLVECTLKHAFEGASMKNLVLIISGSPVSLHVSYDLSLSOLF	240		
QY	241	RNPORPVSNTLEKGLFARIEKFLSPOLIAEEFLKTFPSGQPIPARPASGONSI	300		
DB	241	RNPORPVSNTLEKGLFARIEKFLSPOLIAEEFLKTFPSGQPIPARPASGONSI	300		
QY	301	SVMPAKITTKRAKYGPVLYAKYGGDKLHEKFLDKNHQAOHTEPKRANGGEERKISE	360		
DB	301	SVMPAKITTKRAKYGPVLYAKYGGDKLHEKFLDKNHQAOHTEPKRANGGEERKISE	360		
QY	361	EAARRRLEETFEKKKQDQIISLMKAEQKKROEKERLEINRNARQGRNVLSAGSGGE	420		

```

Db 361 ENAKRRLEFTEKREKKOD-IRFLAKAOKROKORLEKRLINAKDOGRNLRAGSGGE 419
Oy 421 VKAPFLSGGTIAPSSFSRSGOYEHYHAIFDMOQOARADNAKMKREYIGRLPEROG 480
Db 420 VKASFFIGGAVSPSPGQYEHYHAIFDMOQLRADNARKKGGYGHMLPEROGK 479
Oy 481 QLAVERAKOYEELQORREKMONKARABAGHMYVLARLQIRLONFNEROQIKAKLGEKK 540
Db 480 HLAVERANOVYEELQORREKMONKARABAGHMYVLARLQIRLONFNEROQIKAKLGEKK 539
Oy 541 EANSSEOGESSEADMRKRIEESILKAHANAARAVLKEQLEKREKREKVEEKKVMEHLYA 600
Db 540 EADQTKQEAETEEDMRKRIEESILKAQTNARAVALKEQLEKREKREKVEEKKVMEHLYA 599
Oy 601 KGVSSVSPPLGHEGSGSPKQOMRVSYSALKVEGVDSLTDPRETSSEOKTNN 660
Db 600 R-VASSDVPRLPELLETGSGSPKQOYKPVYSTALKVEGLDGLSDIYOE--EENEKSN 656
Oy 661 AISKREILRLNENLKAODEKGMQNLSDTEINVEDAKHEKESVSDRKKEWEG 720
Db 657 AISKREILRLNENLKAODEKGMQNLSDTEINVEDAKHEKESVSDRKKEWEG 716
Oy 721 QLVIPDLDELDTSTFSTERRHVGVEYIKLGPNGSPRRANGSPDTSVLIIGAELEQL 778
Db 717 QLVIPDLAVLDTSTFSTERRHVGVEYIKLGPNGSPRRANGSPDTSVLIIGAELEQL 774

```

```

RESULT 2
17885
serine/threonine-specific protein kinase (EC 2.7.1.-) SK2 - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Sep-1999
C:Accession: I78885
R:Levedekou, E.N.; He, M.; Bapclist, E.W.; Craven, R.J.; Cance, W.G.; Welch, P.L.; Simic
O:Oncogene 9, 1977-1988, 1994
A:Title: Two novel human serine/threonine kinases with homologies to the cell cycle regu
A:Reference number: I58396; MIMD:9426868
A:Accession: I78885
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-841 <RES>
A:Cross-references: GB:L20321; NID:g348244; PTDN:AAA36658.1; PID:g348245
C:Gene: GDB:SK2
A:Cross-references: GDB:374125
A:Map position: 3p21.1-3p21.1
C:Superfamily: human serine/threonine-specific protein kinase SK2; protein kinase homo
C:Keywords: phosphotransferase
F:4-261/Domain: protein kinase homology <KIN>

```

```

Query Match 13.6% Score 847.5; DB 1; Length 841;
Best Local Similarity 28.1%; Pred. No. 3, 1e-19;
Matches 263; Conservative 160; Mismatches 316; Indels 197; Gaps 34;
Oy 4 YVRLOKIGESFGKALIVKSTEDGRQYVKEINISMSSKRESEKREYAVLANKKHNT 63
Db 6 YCYLRVVGKSGYEVTYLRHRDQKQYIKLKLNRASSRERRARAEQALISQKHPNI 65
Oy 64 VOYRESEF-ENGSLYIVMDYEGGDLFKRINAKGVLFQEDQILDMFVOICLAKHVDR 122
Db 66 VTYRESEWEGDGLLYIVMGCEGGDLYRKLEQKQLLPENOVVEFVQIMAAQLYLHEK 125
Oy 123 KILHRDIKSONIFLTKDGTVALDQFGIARVLNSTVELATGCTGTPYLSPELCEKPPNN 182
Db 126 HILHRDIKTONVFLTRINIKVGLDGIARVLEWHCMASLTGTPYMSPELFSKPPNY 185
Oy 183 KSDIHALGCVYELCTLKAHFEAGSMKLVKLTISGSPVSLHYSYDLSISQDLFRN 242
Db 186 KSDVWALGCCVYEMATLKAHFNKMDKMSLVYRIIEBKLPAMRDSPELAEILIRMLSKR 245
Oy 243 PRDRPSVNSILEKFIKRIEKLFLQLIAEFCLTFSKFSQPIPAKRPPASGONSISV 302

```

```

Db 246 PERRSVRSILKQPIYIKQISFLE-----ATIKSKNNT-- 281
Oy 303 MPQAKITRPAKANYGIPLAVKRYGKLLHEKKPL-----QKHQAOTPEKRVNIGE 353
Db 282 -----KNGDS--QSKPFYTVSGAESNHEVTHPQ----- 310
Oy 354 ERKISEEAKRRRL--EFIEKREKQDQIT--SLMKAQMKRO-----EKERLERIN 402
Db 311 ----LSSEGSQYIIGEEKCLSOEKPRASGLKSPASLKHTCKODISNTTLETATISVN 366
Oy 403 RAREGQNRVLSAGSGGEVKAFLFGSGGTIAPSSFSRSGOYEHYHAIFDMOQOARDE 462
Db 367 -----IDLPAKRDSVSGFQEN--QPRILASNLNGLGICSI-SQVEEEMQDNT 415
Oy 463 AK-----WKREYGRGLPEROG-QLAVRAKQVEEFLQORREKMONKARABAGM 511
Db 416 KSSAQEPNLIIPWSSDIYTGKNEPVKPLQLEQKPKQDSL-----ALSPLECSGTI 470
Oy 512 VYLARLQIRLONFNEROQIKAKLGEKKEANHSEGGSEEDMRKRIE-SLKAHANA 570
Db 471 LAHSNL--RLGSSDSPASASRAVAGITGVCHADQDVAGECTIEKGRTHPDLPHNSG 527
Oy 571 RAAVLKEQLEKRRKAYERERKWEHLVAKVKSQVSP--LQGHETGSGSPSKQOMS 628
Db 528 SEPSLSRORRORRREOTHR--GEKROVRRLDLPAPQESPPRLPSPHYGK----- 576
Oy 629 VISVTSALKEVG-----VDSLTDTRETSEEMOKTNNALSKREILRLNENLKAODEK 683
Db 577 -VDVSTQKEAENORRVYTGVSSSR--SEMSSSKDRPLSARE-RRRLKO--SOEEMS 629
Oy 684 GQONLSDFEINVEDAKEHEK-----SVSSDRKKEWEGQVLVPLDELTD 732
Db 630 SSGSPVRSKSLSVAPBGRPOEDQPLPARLSDCVTOGRQIHC-----LSEDELSS 684
Oy 733 TSESTERRHVGVEYIKLGPNGSPRRANGSPDTSVLIIGAELEQIOTELL-----E 784
Db 685 TSTSDKSDQDYE-----QKQOTNEI--NALVOLTOVLKIDKSESE 725
Oy 785 NNTINSEIS--PEGKYRPLITGEKVOCSISEIN--PSAIVDSFVETKSESESA-- 836
Db 726 DVPVNPVSEFKLRKYDITLHGKVAEEAEIHEKELPSALM--PGSEKIRRLVEYLR 783
Oy 837 -----SPQMSLKLEGNLEPDLLETI-LQEDSG 864
Db 784 TVIRIGLGVQLIEQVYDLLEDEDFDREVRLREHMG 819

```

```

RESULT 3
JC7122
protein kinase (EC 2.7.1.37) 2 - mouse (strain balb/c)
N:Alternate names: serine (threonine) protein kinase
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001
C:Accession: JC7122
R:Hayashii, K.; Igarashi, H.; Ogawa, M.; Sakaguchi, N.
O:Biochem. Biophys. Res. Commun. 264, 449-456, 1999
A:Title: Activity and substrate specificity of the murine SK2 serine/threonine kinas
A:Reference number: JC7122; MIMD:20001940; PMID:10529384
A:Accession: JC7122
A:Molecule type: mRNA
A:Status: Preliminary
A:Residues: 1-792 <HAY>
A:Cross-references: GB:AJ223071; NID:g4138208; PID:g4138209
C:Gene: SK2L
C:Superfamily: human serine/threonine-specific protein kinase SK2; protein kinase ho
C:Keywords: phosphotransferase

```

```

Query Match 13.4% Score 836; DB 2; Length 792;
Best Local Similarity 29.0%; Pred. No. 6, 6e-19;
Matches 238; Conservative 129; Mismatches 269; Indels 186; Gaps 22;

```



```

OY 240 KRNPDRPSVNS-----ILEKFIARKIEKELISPOLIAEFLCLTKFSKFSQIPIAKRP 293
DB 223 QGNVKEEPRYSAKRMASEILKHPLDPYVEOY-----RP 256
OY 294 ASGNSISVMPAOKITRPAKYGIPLAYKKYGDKLHEKKPLQKHNQAHQPREKRVNIGE 353
DB 257 TLSAASI-----TPREKPLNSRE 273
OY 354 ERRIISEBAARKRLFEIEKEKKOKDOILISMAEQMKROEKERLERINAREQGMRYVL 413
DB 274 GRSMAESON-----SNSSSEKDMFY-----VSDKNIRYVV 304
OY 414 SAGSGEYKAPFLGSGTIAFSSSSNGOYENHAFEDMOOQRAEDN-EAKMKREITGR 472
DB 305 PSNGN-----KYLETDSG-----FVDEDEDLDHVOQSAENGNIQSVSATKRPDGH 348
OY 473 GL-----PERKQGLAYERAKOYEFLOKREKEMOKAKAREGHMYLARIQIRLQNTNE 527
DB 349 GILKPVHSDQRPDYIQRHPRRTINIMVYLKE--EKARENG----- 387
OY 528 RQOIKAKIRGEKKKPAHNSCEQESSEADMRKKIESLKANA--NARAAVLKEOLERRK 584
DB 388 -----SPRMSRSRSPSSVPTQKNNVETPSKIPKIGDI-AHASKTNASTRIPPSKL----- 436
OY 585 EAYEREKKVMWEHLVAKGVASDVSPRLQHEGSGSPSKQOMKSVYISVALKEGVDS 644
DB 437 -----ASD-----SARTPGSPRPNNHNPVI----- 456
OY 645 LTRRETSEEMOKTNNAISSKREILRLNENLKAQDEKQMONLSDTFEIVNEDAKENE 704
DB 457 -----DSSPKLKPRNDRTSP-----SPAANHE 478
OY 705 KEKSVSDRKKWEAGQGLVIRDELTLDFSSTTERHTVGEVILKIGPNSPRRAGKSPRT 764
DB 479 AEEMASVYKR-----QRRPLRPRT----- 499
OY 765 DSVLKIGELALQOTELLENTTIRSEISPEGEKYKPLITGEKKVOCISHEINPAIYDS 824
DB 500 -----SLIHGSRQLGAD-----ISMAAKETAKLHPVSPSEEN--SHO--SRVNAS 544
OY 825 PVETKSP-----PSEASPOMSLKEGN-LEERDDLETLIIOEPSGTNKE-- 869
DB 545 PVST-TPEPKRTSVGSAKQMSSESNISISSLSQAFELCDASPRYIDMTENTTPDDHR 603
OY 870 -----SLPCTITDVWISSEKET--KETOSADRTIOENES-----EDGVSTV 911
DB 604 RSCHSEVYSFPRDISSEMTIRDBHSTSMRLTEIPDSVSGOVQWTIANHOREEGSCPTV 663
OY 912 ---DOLSDIH-IEPGTNDQSHKCDYKSVQPERFFHAKVHNSPHNLVPOVOVOCSPRE 967
DB 664 LKDISPRALQYEPNTSOHQHD-----DKFTVKEFVSVPGRAPRLHYEP 710
OY 968 SFARSHSHLPKKNKNSLLIGSLGPLDANPKMLRCSLPDISKL-----FRLMDV 1022
DB 711 SHOVNSHDNKTYSVNSNL-----EKNNSHHPRVYDVYHVRHSFRVGSQ 762
OY 1023 PTVG--DYRONLE-----IDIKDE-NIKGPRSDSDIYFEETDDLOLQASMOQLRE 1075
DB 763 PVMSVEVGVOVDMKLVNVRDEMEVRKGAIPSES-----PTTRTISPEPDSRTTERPRE 819
OY 1076 -QPGBEISEEESVLAKNSDVEPTANGTD--VADEDDNPS 1111
DB 820 PDPITNYS-ETKSFNSCSDSSPAETRTNVSFVPEETPT 857

```

RESULT 6
 G01452
 NIMA-like protein kinase 1 - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
 C:Accession: G01452
 R:lin. K.P.
 submitted to the EMBL Data Library, June 1994

```

A:Reference number: G07172
A:Accession: G01452
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-445 <LUX>
A:Cross-references: EMBL:U11050; NID:9507874; PID:AAA19558.1; PID:9507875
C:Genetics:
A:Gene: NKL1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: ATP
F:6-271/Domain: protein kinase homology <KIN>
F:14-22/Region: protein kinase ATP-binding motif

Query Match      8.8%  Score 550; DB 2; Length 445;
Best Local Similarity 33.3%  Pred No. 1.8e-10;
Matches 145; Conservative 68; Mismatches 156; Indels 66; Gaps 13;

OY 2 EKYVRLQKIGESFGKALVSTEDGROYVYKEINISMSSKEREERREAVLANMKRP 61
DB 6 EDEYVLYTIGTSYGRQKIRKSDGKILVWKELDYSGMTEAEKOMLVSEVNLRLKHP 65
OY 62 NIVQYRESE--EENGSLYVMDYCEGDLFRINAQGV-----LFOBDQILDMFVOICLA 115
DB 66 NIVRYDRIIDRTNTLYIVMEYCEGDLASVIT--KGTKEROYLDEDFLRVWTLTLTA 123
OY 116 LKHVHDRK-----ILHRDKISONIFLTRKDGTVQAGFGLARVLSVTELARTGTPRYVL 170
DB 124 LKCHRRSDGHTVLRDLKPAVNFELDGKQVWKLGDFGLARILNHDTSFKAFTVGTPTYVM 183
OY 171 SPEICENKRYNKSQDIMALGCVLYELCTLKHAFEGASMKNLVLTIGSPSPVSLHYSD 230
DB 184 SPQNMNRMSVYNSKDOIWSIGLIVELCALMPFAFSQELAGKIRGKRRIRPYRTSDE 243
OY 231 LRSLVQLFKRNPDRPDRPSVNSITLKEGFIARKIERFLSPOLIAEFLCLTKFSKFSQPIA 290
DB 244 LNEITIRMLNLKDYHNFVSVEILLENPLA-----DLVAD-QRNLERGRQ---L 290
OY 291 KRPASQNSISVMPAOKITRPAKYGIPLAYKKYGDKLHEKKPLQKHNQAHQPREKRVN 350
DB 291 GEPEKSDSSPVLSLKL-----KEIOLDEREALAREERLE 328
OY 351 TGBE---RRKISEE-AARKRRL--EFIEKKR-----OKQDILISMAEQMKR----- 392
DB 329 OKQDELCVBERLAEDKLAENLILKNYSILKERKFLSLASNPDLNLPSSVIKKVHFSQ 388
OY 393 QEKERLERINRAREQ 407
DB 389 ESKENIMRSENSESQ 403

```

RESULT 7
 T29771
 hypothetical protein ZC581.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
 C:Accession: T29771
 R:Waterston, B.; Gattung, S.; Lee, T.T.
 submitted to the EMBL Data Library, May 1997
 A:Description: The sequence of C. elegans cosmid ZC581.
 A:Reference number: Z20682
 A:Accession: T29771
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-357 <MAT>
 A:Cross-references: EMBL:AF003134; PIDN:AA854139.1; GSPDB:GN00019; CESP:ZC581.1
 A:Experimental source: strain Bristol N2; clone ZC581
 C:Genetics:
 A:Gene: ZC581.1
 A:Map position: 1
 A:Introns: 31/3; 81/1; 120/3; 186/1; 226/3; 260/1; 288/3
 C:Superfamily: Ca2+/calmodulin-dependent protein kinase I; protein kinase homology

Query Match 8.7%; Score 540.5; DB 2; Length 357;
 Best Local Similarity 37.6%; Pred. No. 2.8e-10;
 Matches 121; Conservative 60; Mismatches 118; Indels 23; Gaps 7;

1 MEKYVRLQKIGEGSGFKALIVKSTEDGROYVYKEINISRMSSKREESREAVAVLANM 58
 1 MONTYKAVVGRGAFVCMWLCRGKNDASHQKVIKILNTHGMEKEENSIOSEVWLTKKV 60
 DB 1 MONTYKAVVGRGAFVCMWLCRGKNDASHQKVIKILNTHGMEKEENSIOSEVWLTKKV 60

59 KHPNIVYRESFEENGSLIYMDYCEGDLFKRINAKOV-----LFOEQDQILDMFVQ 111
 61 QHPILIGYIDFIMDNOGLIYMOYAEGLERLINDORAIKOSNMREYPERKTVLDYEFQ 120
 DB 61 QHPILIGYIDFIMDNOGLIYMOYAEGLERLINDORAIKOSNMREYPERKTVLDYEFQ 120

112 ICLAKHYHDKKILHROIKSONIFETKDGTV-OLGDEGARVNSTVELARFCIGTPYVL 170
 121 ILIALHMHOKVIVHROIKSONILMNRKTYLKLSDPGISKEL-GRKSAASTVIGTPYVL 179
 DB 121 ILIALHMHOKVIVHROIKSONILMNRKTYLKLSDPGISKEL-GRKSAASTVIGTPYVL 179

171 SPICENKRYNNKSDIYALGCVLYELCTKHAEPAGSMKLVKLTSGSPVPSLHYSVD 230
 180 SPECISRYNNKSDMSGLCVLYELQLERAFDGENPRAIVKTIKRSKONPLGDHVSMD 239
 DB 180 SPECISRYNNKSDMSGLCVLYELQLERAFDGENPRAIVKTIKRSKONPLGDHVSMD 239

231 LRSIVSOLFKRNPDRPSVNSTLEKGIATKIEKFLSPOLAEFCLTKFSKGSQPIPA 290
 240 VKMLVNLKTKHTDRKRPVSOGLS-----DPLVPLYLISHDL--GRLEPPPTK 288
 DB 240 VKMLVNLKTKHTDRKRPVSOGLS-----DPLVPLYLISHDL--GRLEPPPTK 288

291 KRPAAGONS-ISVMPAKITKP 311
 289 RKPSALSRLRTYPTOSTLRP 310
 DB 289 RKPSALSRLRTYPTOSTLRP 310

RESULT 8
 A43734
 probable protein kinase nima (EC 2.7.1.-) - *Emmericella nidulans*
 C:Species: *Emmericella nidulans*, *Aspergillus nidulans*
 C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 24-Sep-1999
 C:Accession: A43734; S28786
 R:Osmani, S.A.; Pu, R.T.; Morris, N.R.
 Cell 53; 237-244, 1988
 A:Title: Mitotic induction and maintenance by overexpression of a G2-specific gene that
 A:Reference number: A43734; MUID:88194523
 A:Accession: A43734
 A:Accession type: DNA
 A:Residues: 1699 <OSM>
 A:Cross-references: GB:M20249; NID:9168065; PIDN:ANA3316.1; PID:9168066
 C:Genetics:
 A:Gene: nima
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
 C:Keywords: ATP; autophosphorylation; nucleus; phosphoprotein; phosphotransferase; serin
 F:9-295/Domain: protein kinase homology <KIN>
 F:17-25/Region: protein kinase ATP-binding motif

Query Match 8.5%; Score 530.5; DB 2; Length 699;
 Best Local Similarity 27.8%; Pred. No. 1e-09; Indels 139; Gaps 18;
 Matches 168; Conservative 92; Mismatches 206;

2 EKYVRLQKIGEGSGFKALIVKSTEDGROYVYKEINISRMSSKREESREAVAVLANMKHP 61
 9 DRYEVLKIGGSGFGIIRKVRKSDGFLCRKEINIKSTEREDLAEFNILSLRHP 68
 DB 9 DRYEVLKIGGSGFGIIRKVRKSDGFLCRKEINIKSTEREDLAEFNILSLRHP 68

62 NIVYVRESFEENGSLIYMDYCEGDL-----FKRINAKOVLLDEQDQILDMFVQIC 113
 69 NIVYVRESFEENGSLIYMDYCEGDL-----FKRINAKOVLLDEQDQILDMFVQIC 124
 DB 69 NIVYVRESFEENGSLIYMDYCEGDL-----FKRINAKOVLLDEQDQILDMFVQIC 124

114 LKLVH-----DKRLHDKIKSONIFETKDGTVOLGDEGARVNSTVELARFCIGTPYVL 146
 125 TALYRCHVGTDAEYGSNMLGPAPKPSGLKMGKQAMTILHRDLPENIFGSDNIVYKLD 184
 DB 125 TALYRCHVGTDAEYGSNMLGPAPKPSGLKMGKQAMTILHRDLPENIFGSDNIVYKLD 184

147 FGIAVRLVSTVELARFCIGTPYLSPEICENKPYNNKSDIYALGCVLYELCTKHAFAENG 206
 185 FGISKLHSH-DFASTYVGTPTPYMSPICAEKTYLRSIDIAVAGCIINTELCOSEPFPNR 243
 DB 185 FGISKLHSH-DFASTYVGTPTPYMSPICAEKTYLRSIDIAVAGCIINTELCOSEPFPNR 243

207 SMNLVWLKLTSGSPVPSLHYSVDLRSIVSOLFKRNPDRPSVNSTLEKGIATKIEKFL 266

244 THIQVOKIRGCKPAPLPDSEELKNVIAQLRVDPDRPTAVLINTPYI----- 295
 267 SPOLIAEFCCTKTSKGSOPIPAKRPASGONSISVMPAKITKPAAKYGIPLAKYKID 326
 DB 267 SPOLIAEFCCTKTSKGSOPIPAKRPASGONSISVMPAKITKPAAKYGIPLAKYKID 326

296 -RLMRREVELNINS-----RAARKREATOKAKDVEQAFK----- 331
 327 KILHEKKPLOKHOAHOPKERVNTGERRKISEEARRRLTEFEKKOKQDIISLMK 386
 DB 327 KILHEKKPLOKHOAHOPKERVNTGERRKISEEARRRLTEFEKKOKQDIISLMK 386

332 -----LEKQOIRS-----ELENSTIREVEVAKLEIDROVONELDKLRKFE 375
 387 AQOMK--QEKERLERINAREOGWVLISAGSGEYKAPFLSGGTIAPSSFSRGOYE 444
 DB 387 AQOMK--QEKERLERINAREOGWVLISAGSGEYKAPFLSGGTIAPSSFSRGOYE 444

376 CEVODVABEYVOKRANNTYREDA-SLRSGHSSQK-----SSNSSDSDPSSSTO-- 425
 445 HYHALFDQMOORADNPAKKRELIGRGDPERQKGLAVERAKOYEERLQKREAMONK 504
 DB 445 HYHALFDQMOORADNPAKKRELIGRGDPERQKGLAVERAKOYEERLQKREAMONK 504

426 -----ISQLSLSPYNNKAK-----LPKK-SKTPPTRSKTV-----DSPMDI 463
 505 ARAEGHMYLARIKRILOFNERNQIKAKLIGKFKRANSEGOBSEADARRKRIEEL 564
 DB 464 QMAEPSPISIALS-----SLSPKTSNAYS-GKNIFAGEKKRPREFPT 506

565 KAHAN 569
 507 LAYS D 511
 DB 507 LAYS D 511

RESULT 9
 T37970
 probable G2-specific protein kinase (EC 2.7.1.-) - fission yeast (*Schizosaccharomyces*
 C:Species: *Schizosaccharomyces pombe*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
 C:Accession: T37970
 R:Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 Submitted to the EMBL Data Library, September 1997
 A:Reference number: Z21758
 A:Accession: T37970
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Accession type: DNA
 A:Residues: 1722 <RAD>
 A:Cross-references: EMBL:Z98975; PIDN:CA811653.1; GSPDB:GN00066; SPDB:SPAC19E9.02
 A:Experimental source: strain 972h-; cosmid c19E9
 C:Genetics:
 A:Gene: SPDB:SPAC19E9.02
 A:Map position: 1
 A:Insertions: 20/3; 28/3
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: phosphotransferase; protein kinase

Query Match 8.2%; Score 514.5; DB 2; Length 722;
 Best Local Similarity 24.2%; Pred. No. 3.2e-09; Indels 257; Gaps 28;
 Matches 199; Conservative 116; Mismatches 250;

1 MEKYVRLQKIGEGSGFKALIVKSTEDGROYVYKEINISRMSSKREESREAVAVLANMKH 60
 1 MEKYVRLQKIGEGSGFGIIRKVRKSDGFLCRKEINIKSTEREDLAEFNILSLRHP 60
 DB 1 MEKYVRLQKIGEGSGFGIIRKVRKSDGFLCRKEINIKSTEREDLAEFNILSLRHP 60

61 PNIVYVRESFEENGSLIYMDYCEGDLFKRINAKOV-----GVLPQDQILDMFVQIC 115
 61 PNIVYVRESFEENGSLIYMDYCEGDLFKRINAKOV-----GVLPQDQILDMFVQIC 119
 DB 61 PNIVYVRESFEENGSLIYMDYCEGDLFKRINAKOV-----GVLPQDQILDMFVQIC 119

116 LKLVH-----DKRLHDKIKSONIFETKDGTVOLGDEGARVNSTVELARFCIGTPYVL 179
 120 IYRCHVGENAPACDSOMPREFHKGOSVLRHDKIPAFIFDENNSVKLDDFGISLMDPT 179
 DB 120 IYRCHVGENAPACDSOMPREFHKGOSVLRHDKIPAFIFDENNSVKLDDFGISLMDPT 179

157 VELARFCIGTPYLSPEICENKPYNNKSDIYALGCVLYELCTKHAFAENGSMKLVKTI 216
 180 RVFTQSYVGTPTPYMSPICAEKTYLRSIDIAVAGCIINTELCOSEPFPNR 239
 DB 180 RVFTQSYVGTPTPYMSPICAEKTYLRSIDIAVAGCIINTELCOSEPFPNR 239

217 SGSPFPVPSLHYSVDLRSIVSOLFKRNPDRPSVNSTLEKGIATKIEKFLSPOLAEFC 276

Db 240 OGNLSCMDHHYSDVFFILLRHCLFVNSDLRPTTYQLLRSPILSDIRKSLSESRVLEQSD 299
 QY 277 LKTFKFSQSPITPAKRPAAGONSISVMPAKITKPAKAVGIPLAYKKYGDMLKHEKKPQ 336
 Db 300 L-----K 303
 QY 337 KHKHAKHPEKRVNTGEPERKRISEAKRKRLFEIEKEKKQKQDQIISLKAQMKQKE 396
 Db 304 KHKMLQL-ENDDQFREQRL-----SKRESELEVLASR-----LAQREILRELE 349
 QY 397 FLERINRAREOGKRVNLSAGSGEVKAPLGSCTTAPSSFSRGQYEHYHAIPDOMOO 456
 Db 350 KOLRMDARYO-----RHMQTVVNSMOM 373
 QY 457 R---AEDNEAKMKREIYRGCLPEROKGLAVE---RAKQ-----VEEFLOR--- 496
 Db 374 RYTSFVDHNEQ-----PESSTAEKMFVDCCTTEASQSPLLHLPKLGISKPLQTLSC 422
 QY 497 -----KREAMQNKARAEQ-H-----MYVLARLQRLQNFNEROOIKAKLAGE 538
 Db 423 PGFTLTTOOPLIKRPTLRKELSRALHTTATLMKYRAMSSLRTPIDKGOI-TSLQOK 481
 QY 539 KKEAHNSGQEGSEADMRKRIEISLKAHANARAVALKEQLEKREKREKREKVEEHL 598
 Db 482 NGTSN-----QVADCMNK-----LLHTSLDGKTLSPSELCKNFSD--- 516
 QY 599 VAKGVKSSDVSPPLGQHEGTGSPSKQOMRSVISTALKFVGVDSSLTPTRETS----- 652
 Db 517 -GEGLPNRKVS-----KLSVESDEFAVAGASSG-ESVPTDSTLTDTKSKSVYVHP 564
 QY 653 -----EEMOKTN-----NAISSKREILRLN-----ENLKAODEKGMON 687
 Db 565 SPOSILYVKELKLIIRDSVSKASKATLIGALSLASPIYVAHEKRIARENE-MDG 622
 QY 688 LSDTFEIVHEDAKHEKEKSVSDRKKEWAGQGLVPLDEL 729
 Db 623 NFKTKINQHPD-----EYVLRTPKKIQLEQKRSVPKQL 658

 RESULT 10
 S59359
 N:Alternate names: yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 C:Date: 30-Nov-1995 #sequence_revision 01-Mar-1996 #text_change 24-Sep-1999
 C:Accession: S59359; S69565
 R:Longline, M.S.: Pringle, J.R.
 submitted to the EMBL Data Library, August 1995
 A:Reference number: S59359
 A:Accession: S59359
 A:Molecule type: DNA
 A:Residues: 1-1142 <ION>
 A:Cross-references: EMBL:U03140; NID:g992650; PIDN:AAA7513.1; PID:g992651
 R:Dieckrich, F.S.
 submitted to the EMBL Data Library, August 1995
 A:Description: The sequence of S. cerevisiae cosmid 8166, 9787, 9717, and lambda 3073.
 A:Reference number: S69565
 A:Accession: S69565
 A:Molecule type: DNA
 A:Residues: 1-1142 <DIE>
 A:Cross-references: EMBL:U03057; NID:g927764; PIDN:AA64949.1; PID:g927777; MIPS:YDR507C
 C:Genetics:
 A:Gene: SGD:GIN4
 A:Cross-references: SGD:S0002915; MIPS:YDR507C
 A:Map position: 4R
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
 C:Keywords: Arp
 F:17-289/Domain: protein kinase homology <KIN>
 F:25-33/Region: protein kinase Arp-binding motif

Query Match 8.0%; Score 502.5; DB 2; Length 1142;
 Best Local Similarity 19.8%; Pred. No. 1.1e-08;

Matches 252; Conservative 215; Mismatches 475; Indels 331; Gaps 42;
 QY 8 OKIGESFGKALIVKSTEDGRQYVKEI-----NLSRMSKREERES-----RREV 52
 Db 23 ETGLGSLTGQVLARNGSTGQEAAYKISKAVNFNGVNSGTSYSGTTPALPGIEREI 82
 QY 53 AVLANKHRIYQYRSSEFENSITYVNDYCGGLFRIKINAGVLPQEDQLDHWVOI 112
 Db 83 IIMKLHNPVNLIDYVETNTDLVLETAAGKGLFNL-VEKGL-PEHEALRRFRQI 140
 QY 113 CLAKVHDKILHRDKISQNIETKDGTVOLGDFIARVLNSTVELARQIGTPYIPLS 172
 Db 141 IIGVSYCHALGIYHRDLKPENLILLHKYNIKIAIDGMA-ALETGKLETSCGSPHAP 199
 QY 173 EIGENRPYNN-KSDIWLGCVLIELCTLKHAQ-EAGSMKLVKITSGSP-PVSLHYS 228
 Db 200 EIVSGIPIYOGFASDVWSCVILFALLTGLRPEDEEDGNIRLLKVKRGFEFMSDDEIS 259
 QY 229 YDLRSIVSQLFKHNPDRSVNSILEKGFIAK-----RIEFLSPOLIAEE 274
 Db 260 REAQDLIRKILYVDPERRIKTRDILKHPLOKYPISIRDSKIRGLPREDYVLPFSNS 319
 QY 275 FCLKTFESK-----FGSOPITPAKRPAAGONSISVMPAKITKPAK-----YGIPLAYK 323
 Db 320 SIDATILQNLVILMHGRDEGIK-----EKLREPANMEKTLVALLYRFK 365
 QY 324 YGDKKILHEKKPILOKHQAQO-----TPEKRYNTGEERRK-----ISEPARRRLEFIEKEK 375
 Db 366 DTQKELIKQOQVKKRQSTISSVSFSPKYSTTPQRRRRESLIVTSRKKPIEF-NKFT 424
 QY 376 KQKQDQIISLKAQMKQKEKRELKINRAREOGKRVNLSAGSGEVKAPLGSCTTAP 435
 Db 425 ASSASSSNLTTPGSSKR-----LSKNFSSKKLSTIVWSSPTPA 464
 QY 436 SFSSRQYEHYHAIPDOMOOQRAEDNEAKMKREIYGLP-----EROKGLAVE 485
 Db 465 SRNRKAS-----VINENKQKRASIFSTYKKNRSSRISIKRMSLIPSKRRSVYTKMST 519
 QY 486 RAKQVE-----EFLQK-REAMQNKARAEQHYVLARLQRLQNFNEROOIKAKLAGEK 540
 Db 520 YARLAEDDDWEYIEKETKRTSSNFATLIDEIFEYKYQIR-----KKEBELERKVRBAK 574
 QY 541 EAHNSGQEGSEADMRKRIEISLKAHANARAVALKEQLEKREKREKREKVEEHLVA 600
 Db 575 -----ARELEKRRKQKEKE-----RAKRLKEEDLKKQEDL-----KKQIE 613
 QY 601 KGVKSSDVSPPLGQHEGTGSPSKQOMRSVISTALKFVGVDSSLTPTRETSPEMOKTN 660
 Db 614 --IDIDLQGLSKH-----KEKIDGNIRISAPMEHEENIN 650
 QY 661 AISKRE-ILRLNENLKAQODEKGMONLSDFEIVHEDAKHEKEKSVSDRKKEAG 719
 Db 651 HLEVYDNLIRKRNFSQIR-----KEKIDGNIRISAPMEHEENIN 650
 QY 720 GQIVLPDELTLDTSTSTERTVGEVYIKLGPNSPRAMKSPDTSYLKIGEAELQLO 779
 Db 671 -----PVSRLDPCIMS-----SPEEEV----- 688
 QY 780 TELLENTTIRSETISPEGEYKPLITGEKKVOCI--SHEINSAIVDSVPETKSPSEAS 837
 Db 689 -----SPEPKRTENERLTTEKILLETIRSKSLGSSFNIDELKLSKMEYPSII 738
 QY 838 PQMSLKLEGNLEPPDLEIILQEPSTGNKDESLCTTIDVWISE----- 882
 Db 739 APQRLSEERVSDNSNGYESLILPDKNGVSQLKOSTATAPVSDGLRKISERIVPQFT 798
 QY 883 EKTETKQSDRTTQIENFVSDGVSTVVDLSIHIPGNDGSHKCOVDKSVQDEP 942
 Db 799 RKRHRFSESKRRLSVLSMSTKESFTNLVDILKQNDL--VNQOSQRIPTPSADDER 856
 QY 943 FKHVYSEHLNLVPOVQSQCPESFAPRSHSLPFPKKNKNSLLGLSTGLFDANPK 1002
 Db 857 LFTVNE-----AETGNSNDRRLVDVGDSITIKDSALKLNFADRFNGSNEAK 906


```

OY 1003 MLRTCSLPDLSKLEFRTLMADVPVGVKODNMLDEIKDENIKEGCS---DSEDIYFEETD 1059
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 907 QTDNHLILP-----PLNGD-----NELRKNOSQSGOAHFKIKSMIPESGS 949
OY 1060 TLOELQASMQLLREQGEYESEESVLKN---SDVEPTANGTDVADDDNPSSESAL 1116
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 950 SSIHTEKEENEKEEKKP-EQHKOEEOEKREKVDDMEPLKSVQKIREKNAGSQA-- 1006
OY 1117 NEMWSDNSDGIASECDSVFNHLELRLHLEOEMGEPEFEYEKIKAIHEDEDENI 1176
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 1007 -----KDSK-----DHLKEHKODKNTALGNSFFRKFSK-----SSDKTM 1042
OY 1177 ETCISKYVNIICN 1189
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 1043 ELVAKISAKOLF 1055

RESULT 11
T14157
serine/threonine protein kinase - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14157
R:Pytowski, B.; Hicklin, D.J.; Kornhaber, G.; Dellaratta, D.V.; Wltte, L.
submitted to the EMBL Data Library, December 1997
A:Reference number: 217894
A:Accession: T14157
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1233 <P>P>
A:Cross-references: EMBL:AF039574; NID:g2773155; PID:g2773156; PIDN:AA896682.1

```

```

Query Match      8.0%; Score 500; DB 2; Length 1233;
Best Local Similarity 20.8%; Pred. No. 1.5e-08;
Matches 272; Conservative 224; Mismatches 508; Indels 304; Gaps 49;

OY 9 KIGSGFGKALIVKSTFGROYVYKKEINISRMSSKERESEREAVANMKHPNIVQYRE 68
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 39 ELGGAGAGKYVKAQNKETNVLAQAVIDRK--SEELDEYVWEIDILASCHPPIVAKLD 96
OY 69 SFEENGSLXTMDYCEGGDLFKRINAKGVLFQ-----EDQILDMFQICLALKHVHR 122
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 97 AFYENNMILIEFCAG-----AVDAVMLELERPLTESQIOVVCQOTLEALNLYLHDN 149
OY 123 KILHDIKSONIFLTGQVQLDQGIARVINSTVELARTICIGTPYVLSPEI--CE--N 177
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 150 KIIHDLKAGNIFLTLDGDLKADGVSAKNTITQRRSPICITPYWMAPEVVMCETSKD 209
OY 178 KPYNKSIDMALGCVLYELCTLKAHFAFGSMKULYKIIISGSPPPS--LHYSYDLASLV 235
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 210 RPYDYKAVWMSIGITLLEMAIEEPHHELMFMRVLKIAKSEPTILAQSKWSSMFKDFL 269
OY 236 SOLFRNPRDRPSVNSILEKGFIAKRIEFLSPOLIAEFCUKTYS--KFGSQ----- 286
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 270 RKLCEKNVDAWRTTSOLLQHPFVVDNKNPKR-ELIAEAKAEVTEVEEGKEDEBEERE 328
OY 287 ---PIAKRPASGONSISVMPAKITKPA-----KYGIPLAYKKYCDKILHEKKPLQ 336
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 329 NALPIPAKRASSDLSIASSEEDKISQWACLESVSRTOSTSEKFSKILINKEPTTD 388
OY 337 KHKQAHQPEKRYN--TGEERK-----ISEFAKKRLLEFTEKKOKDOITSL----- 384
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 389 GPKAVADENASDVNLETGALNDQTVGIHNGRKKRPKLENLPDQDQTVVNSVSEE 448
OY 385 -----MKAEQMKQKE-----RLERINARROGRNVLASGSGEVK 422
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 449 NENNRVLTETNDCLKPEDRKNKEQETLESKLIQSEINDTHTOTMD--LVSOETGEKE 506
OY 423 APF-----LGGGTIAPSSFSKQGEYHNAIFDQMOQ-----OR 457
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 507 ADFOADVNEVGLTKEETQELGDKGTAAKVITSRRSSVGTDEALDQAKAELSKRAAS 566

```

```

OY 458 AEDNEAKKREIYRGGLPEROKQLAVERAKOYEETLQKKREAMONKARAGHVVYLA 517
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 567 GEGDEALVPQTIAKEPTEGPEAGAEPEEGGEVEDQAPQDPAVCEALGOLTSISST 626
OY 518 RQIRLO--NFEROOI-----KAKLGEKKAHNSGEOGSEADMRKKIES-- 563
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 627 TQATLEQPEDEVRQVSSNSIELELRLVATGAELALISSEEAATEDELERKNQKV 686
OY 564 -LKAHANRAA-----VLKEOLEKRRKAEAREKKWHEHLVAKGVSSDVSPPLQ 614
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 687 PKAASQAPASQSEPPHVLPSININSETTENKE-----MALPKPETILPEPE 739
OY 615 HETGG-----SPSKOMRSVTSALKEVGDSLTDRRESEEMOKTNAI- 662
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 740 HEKNDTSGIGSYVENSSGDLNLSISPLSKADSG--SVLSQETFRQKTLKTRKIV 798
OY 663 -----SSKREILRLNE-----NLKAOEDEKGMNLDTEFEINVEDA 700
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 799 DGEVSVTSKIVTSDSKTELEFLRLQRLRELRLQKEEQAQOLNGKLOOQREQIF 858
OY 701 KEHEKESVSSDRKKWEAGQVYIPDELDTDTSTFTEHRTYGEVYIKGPNQSPRRAMG 760
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 859 RREOE--MLSKKROYD--QETENLEKOKOTIERLEQHT----- 895
OY 761 KSPIDSVLKILGEALQLO--TELLEN--TIRSEFISPEGEKKYKPLITGKKVQCSHRI 816
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 896 NRLRDEAKRIKGOEKELSKFQVNLKKNKEVMNEKARRELRLKTRKRELQASQ- 954
OY 817 NPSAIVDPVETKSPSEASQMSIKLEGNL-----EPDLETETLOPSTND 868
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 955 -----HAQOEFGVQKQO-----ELDGSIKTIQOQKALANTRELL-----NNQ 996
OY 869 ESLPCTITDVAWIESEKETE-----TOSADRTITQENEVEDGSVSTQDLSDIHIEG 922
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 997 QLMARARAIWELEERHLDLQKODLKDQYFMQRHQLK----- 1038
OY 923 TNDQSHKCVQDVSQPEPFHKKV-----HSEHLNLYPVQSVQCSPESEF--AFNS 973
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 1039 -----RHKK-----ETEQMQRYNQRLIELEKNROTQERARLPITQSEKTRAMFKSLRI 1090
OY 974 HSHLP---NNKKNNSLLIGLSTGLDANNPKMLTGSILPLSKLFTLMD--VPTVGVY 1028
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 1091 NSTATPQDQKRIKO-----FPAQDEKRNKERRAAGHOKHESQMRDLOLQCEBANY 1140
OY 1029 RODNLEIDIKDENIKRSPDSSEDIYFEETDDLOELQASMQLLRE-----QGEYESE 1083
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 1141 R---ELHQLONEKC-----HLVHEHTQKIKLDEHSEQELKEVREKLRPRKRTLE 1188
OY 1084 EE-ESVLKNSDVEPTANGTVADDDNPSSESALNEM-----HSDNS 1125
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 1189 EEFARKLOEVEFFKMTG---ESECLNPSKQSNISKIFYPIPLHSTGS 1233

RESULT 12
A57177
NIMA-like protein kinase - Emericella nidulans
C:Species: Emericella nidulans; Aspergillus nidulans
C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999
C:Accession: A57177
R:Pu, R.T.; Xu, G.; Wu, L.; Viterla, J.; O'Donnell, K.; Ye, X.S.; Osmanli, S.A.
J. Biol. Chem. 270, 18110-18116, 1995
A:Title: Isolation of a functional homolog of the cell cycle-specific NIMA protein ki
A:Reference number: A57177; MUID:95355415
A:Accession: A57177
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-779 <P>U>
A:Cross-references: GB:LA2573; NID:g1040682; PIDN:AA80145.1; PID:g1040683
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: ATP
F:3-290/Domain: protein kinase homology <Kin>
F:13-21/Region: protein kinase ATP-binding motif

```

Query Match 8.0% Score 499.5; DB 2; Length 779;
 Best Local Similarity 22.0%; Pred. No. 9.9e-09;
 Matches 211; Conservative 129; Mismatches 326; Indels 291; Gaps 31;

2 EKYVRLQIGGSGFKALVSTEDGROYIKENINISMSKREESREVVANLNMKP 61
 5 DKYLELEKIGSGFGIIRKRRADGMLCRLKISYLKSKOKEREDLHAFESILSTLHP 64
 62 NIVQY--RESPENGSLIYVMDYEGGDLFKRI--NAQKGVLFQEDQILDW--FVQICLAL 116
 65 NIVGYHREHLKATQDILHLYMEYCGNDLGRVIRNLKNNQVAEESEFV--WSIFQSLVAL 123
 117 KHHV-----DRKILHRDIKSONIFLTKDGTVOLDFGIA 150
 124 YRCHGVDPPEVGVTLGLSTARPSPSGCMTILHRDLKPEVFLGEDNSYKLGDFGLS 183
 151 RVLSVTEVLAATCIGTPYVLSPEICEKPNPNKSDIMWALGVLYELCTKHAFFEGSMKN 210
 184 KVMQSH-DFASTYVGTPTFVMSPELCAEKYTLKSDIMSLGICITLCAAREPFPNKHQY 242
 211 LVKTIISGSPPVSLHSTYDLRLSVQLFRNRPDRPSVNSLLEKGFIAKRIEKLSPOL 270
 243 LVQKREKRIAPLPVSIGELFATIKDCLRVNDRPRDPTATL----- 285
 271 IAEFCIKTFSKFGSOPIPAKRPASGONSISVPAOKITKPAKYGIPLAYKYGDKLH 330
 286 -----NLPIV-----RLMR 294
 331 EKKPLQKHQAHTPEKRVNTGEBRRKISEAARRLEFIEKKOKDOITSLMKADOM 390
 295 KEKEVEFSRFLRKETELN---KRIRE--LDKSLALETEK-----SSTRAE-- 337
 391 KROEKERLERINRAEOGMNVLVAGSGEVKAPFLGSGGTIAPSSFGQYEHYHAF 450
 338 -----IDASLRREM-----EVKA-----RLRI 354
 451 DOMQOQRAEDNEAKMKREI-----YGRG--LPERQKQOLAVERAQVEEFLRK 497
 355 DRVAQELIESIQCFEEDQVAVAELOGRHGRPFMNSHGQSPFSTATILVSDY---- 410
 498 REAMONKRAAGHNVYLAIRLOIRLQNFNRQOIKAKL-----REKKEAHNSGECBS 551
 411 --NLSSVSGSDGDPSTTDITDITIAESTDGTIKIRPTPRHRAQITYSSAPASVJGT 468
 552 ----EADMRKKITESTLKAHANAAVAVLKEQLERK---KEAYEREKVWEHLVANGV 603
 469 PMDIEMASPSPITIASLSTL-SPRMALTKAPTNTNPMIFGEDEPTSTDKSNMVEVPRETEMI 527
 604 KSSDVS-----PPLQGHETG-----GSPSKQOMRSV 629
 528 DSGESESEALVSPKRTITSSKNPFSTVTRSPSINSQONSNVLPJGLRSKQTLATVR 587
 630 ISVTSALKEVGD-----SLTDRTRESFEMOKTNNAISKRILRLNENKAOEDREG 684
 588 SKTVSGVSSIQHPLRSPAPSLRDKRPPSTRRLSRIPSVTCVGRRLANNINNSNGG 647
 685 MONISDPEFIVHEDAKHEHEKESVSSDRKKMEAGGOIVIPDELDTLQTSFSTERHYG 744
 648 SDAPSTVTSNITVTRGKLRMSSTCD-----ESSFQOQ----- 682
 745 EYIKLPGPGSTRAMGKSPDSDVYKILGEALQLOTELENTTIREISPE-----GE 797
 683 -----NNQPOSLPQAPP--LTKTG-----LMAAKNIRSSSLEVELHQARAGR 723
 798 KYKPLITEKKVQCI-SHEINPSAIVDSPVETKSPESSEASPOMSLKLGNLEEPDD 853
 724 PISAIISNEAKLRAFKREHATIASAVDS-----SSSSSSSSGOSQLPTRPSQPPTD 775

hypothetical protein F20D21.33 (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: B96587
 R:Biologists: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Huizart, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzla
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: B96587
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-200 <SNO>
 A:Cross-references: GB:AB005173; NID:94585993; PIDN:AAD25629.1; GSPDB:GN00141
 C:Genetics:
 A:gene: F20D21.33
 A:Map position: 1

Query Match 7.6% Score 475; DB 2; Length 200;
 Best Local Similarity 43.8%; Pred. No. 1.6e-08;
 Matches 85; Conservative 52; Mismatches 55; Indels 2; Gaps 2;

1 MEKYVRLQIGGSGFKALVSTEDGROYIKENINISMSKREESREVVANLNMKP 60
 1 MEQYEFLEQIGGSGFSLVLRHKKKKYLLKRIAROTQRRAHAEVOLISKMH 60
 61 PNIVQYRESFENGSLIYVMDYEGGDLFKRIAGKGVLFQEDQILDWVQICLAKHV 119
 61 PTVEXKDSWVAKACYVCIVGICGDMQAQIKKNSGNGVHEERKLCQKWLVLMLLEYL 120
 120 HDRKILHRDIKSONIFLTKDGTVOLDFGIRVNSVTEVLAATCIGTPYVLSPEICEKPN 179
 121 HSNHILHRDVKSNIFLTKREDIRGDRGLAKILTSQ-DLTSVVGTPSMCELLADLP 179
 180 YNKSIDNALGCVL 193
 180 YGSKSDIMSLGIFL 193

RESULT 14
 S49313
 protein kinase - slime mold (Dictyostelium discoideum)
 C:Species: Dictyostelium discoideum
 C:Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 24-Sep-1999
 C:Accession: S52076; S49313
 R:Biologists: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Huizart, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzla
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: S52076; MUID:95161460
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1094 <WE2>
 A:Cross-references: EMBL:Z37981; NID:9551445; PIDN:CAA86053.1; PID:9551446
 C:Genetics:
 A:gene: 35/3; 104/1; 166/2
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP
 F/20-281/Domain: protein kinase homology <KIN>
 F/28-36/Region: protein kinase ATP-binding motif

Query Match 7.5% Score 467; DB 2; Length 1094;
 Best Local Similarity 21.5%; Pred. No. 1.3e-07;
 Matches 252; Conservative 208; Mismatches 434; Indels 280; Gaps 50;

Db 513 -AIDNEVGFTKEQKIGKDKTKHKV-----ISDI 543
QY 633 TSLKKEGVYDSSLDTRETSEEMOKTNNAISKR-ETLRRLNENLK-----AOEDEK 684
Db 544 TS---EVGDERPGDTQKSAEQDAEGGAGEARPAQJLTERKATEGPRAHGAEEPR 600
QY 685 MONTSTFEINNHEDAKHEHEKESYSDRKKEAGOLV-----IPDELTLDTSFST 737
Db 601 GERVED-----KOEQOSAVC-----EGEQYTSSESTRATTEPETEDVOY 644
QY 738 TERHTYGEYIKIGPNSPRRAGKSPTDSVLKIIGEAJLOJOTELLNTTIRSEISPEGE 797
Db 645 SESNSTEELERLGYTGAEBQALGSK-----GEATELDLERENA---OELPVKAE 692
QY 798 KYRPLITGKKVOCISHEINPSAIVDSPETKSPFESEASPOMSLKLEGNLEEPDLETE 857
Db 693 POAP-----AASQASEPPVLIPIINIH-ENTENKGE-----GALPKP---ETI 734
QY 858 ILQEP---SGTNKDESLPCTI---TDWMISEKETKETOSADRITIOENEVSESGVST 910
Db 735 LPPEPENGKNDTSGTSTVENSSDLNISISFLSKTKDGSVSLQETRRQKTKTKT 794
QY 911 VQOLSD-IIEPGTNDQSHKCDVYKSVQPEPFHKVYHSEHLNLPQVOSVOCSPESF 969
Db 795 RKFIYDGVESVYT-----SKIVTDSKTEEL--RFLRQELRELRLLOKEQKAQQL 847
QY 970 AFRSHSLPFRKKNKNSLLIGISGLFDANNPKMLRTCSLPDLSKLFRILMDVPTYQDVR 1029
Db 848 -----NGKLOQOREQIIFRPEQMLSKKRQYDQEIENL--EK 882
QY 1030 QDNLEIDETKDNKEGSDSEDIYFEETDTLOLOLQASMEQLRQPGEE-----YSE 1083
Db 883 QOKQTERLEQEHYRNLDEARIRKEQ-EKELSKFO---NMLRNRKEEQEFVOKQO 937
QY 1084 EESVVLKNSDVEPTANGDTVADEDDNPSSSALNEE--WHSNDSDEIASCECDSVFN 1140
Db 938 ELDGALKKIIOOKAKELNIERECINNKOQLLRAREAAIWELIE-----R 982
QY 1141 HLEELRLHLEQEMGEFEVEYEKIKAIHEDEDENIE 1177
Db 983 HLOEKHOLKQOLKDQYFIORHOLKR-HEKETQOMQ 1018

Search completed: May 15, 2002, 07:55:09
Job time: 219 sec

Wed May 15 14:14:05 2002

us-09-783-320-4.rsp

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 07:57:59 ; Search time 33.64 Seconds
(without alignments)
1397.311 Million cell updates/sec

Title: US-09-783-320-4

Sequence: 6243 1 MEKRYVLRKIGSGFGRKIL.....YAKILHLVMDAGYQEDNDE 1214

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3385	54.2	774	1	NEK1_MOUSE
2	847.5	13.6	841	1	SMR2_HUMAN
3	833	13.3	511	1	NEK3_MOUSE
4	790.5	12.7	439	1	NEK3_HUMAN
5	565.5	9.1	431	1	NEKA_TRYPB
6	550	8.8	443	1	NEK2_HUMAN
7	544.5	8.7	443	1	NEK2_MOUSE
8	530.5	8.5	639	1	NIMA_KENT
9	528.5	8.5	431	1	NRKX_TRTIB
10	502.5	8.0	1142	1	GIM_YEAST
11	499.5	8.0	779	1	NIM1_NEUCR
12	468.5	7.5	966	1	ST10_MOUSE
13	455	7.3	968	1	ST10_HUMAN
14	452	7.2	1518	1	KRK1_YEAST
15	437.5	7.0	915	1	KCC4_YEAST
16	432	6.9	435	1	KIN3_YEAST
17	421.5	6.8	1895	1	YLK3_CAEEL
18	393	6.3	1080	1	NRK1_YEAST
19	388	6.2	1062	1	CC7_SCHPO
20	383.5	6.1	705	1	CC5_YEAST
21	382.5	6.1	740	1	K6A3_HUMAN
22	382	6.1	713	1	KP78_HUMAN
23	381	6.1	460	1	CDL1_SCHPO
24	380.5	6.1	685	1	SNK_HUMAN
25	379	6.1	682	1	SNK_MOUSE
26	378	6.1	631	1	SNK_MOUSE
27	378	6.1	733	1	K6A2_HUMAN
28	378	6.1	733	1	K6A2_MOUSE
29	378	6.1	735	1	K6A1_HUMAN
30	378	6.1	735	1	K6A1_MOUSE
31	377.5	6.0	733	1	K6A1_XENLA
32	377.5	6.0	1305	1	GAK_RAT
33	373	6.0	1305	1	GAK_RAT

34	374.5	6.0	982	1	SULU_CAEEL	P46549 caenorhabd
35	374	6.0	350	1	KAP6_HUMAN	P23612 homo sapien
36	374	6.0	615	1	GNK_RAT	O97011 rattus norv
37	374	6.0	745	1	K6A6_HUMAN	O9UK32 homo sapien
38	373.5	6.0	490	1	SPS1_YEAST	P08458 saccharomyc
39	373.5	6.0	646	1	SNK_HUMAN	O914B4 homo sapien
40	371.5	6.0	485	1	K6B2_MOUSE	O921M4 mus musculu
41	371	5.9	350	1	KAPA_BOVIN	O921M4 mus musculu
42	371	5.9	350	1	KAPA_HUMAN	P17612 homo sapien
43	371	5.9	774	1	KEMK_MOUSE	O05512 mus musculu
44	370.5	5.9	444	1	KRAC_DICDI	P54644 dictyosteli
45	369.5	5.9	397	1	KAPA_YEAST	P06244 saccharomyc

ALIGNMENTS

```
RESULT 1
ID      NEK1_MOUSE      STANDARD:      PRT: 774 AA.
AC      P51954;
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DE      Serine/threonine-protein kinase NEK1 (EC 2.7.1.-) (Nima-related
DE      protein kinase 1).
DE      NEK1.
GN      Mus musculus (Mouse).
OS      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RS      SEQUENCE FROM N.A.
RC      TISSUE-Blood;
RA      MEDLINE=93010942; PubMed=1382974;
RA      Letwin K., Mizzen L., Motro B., Ben-David Y., Bernstein A.,
RT      Pawson T.;
RT      NIMA cell cycle regulator and highly expressed in meiotic germ
RT      cells". 11:3521-3531(1992).
RU      EMBL J. 11:3521-3531(1992).
CC      -1- FUNCTION: PHOSPHORYLATES SERINES AND THRONINES, BUT ALSO APPEARS
CC      -1- TO POSSESS TYROSINE KINASE ACTIVITY. IMPLICATED IN THE CONTROL OF
CC      MEIOSIS.
CC      -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC      -1- TISSUE SPECIFICITY: PREDOMINANTLY IN TESTES (GERM CELLS AND
CC      SERTOLI CELLS). LOWER LEVELS IN OVARY (OVOCYTES AND GRANULOSA
CC      CELLS), THYMUS AND LUNG.
CC      -1- DEVELOPMENTAL STAGE: IN FEMALE, EXPRESSED AS FOLLICLES ENTER THE
CC      SECONDARY STAGE UNTIL OVULATION OCCURS. IN THE MALE REPRODUCTIVE
CC      SYSTEM, THE EXPRESSION IS LIMITED TO SPERMATOCYTES AND SPERMATIDS.
CC      -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC      NIMA SUBFAMILY.
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation. Its
CC      content is the property of the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement. (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      EMBL: S45828; AAB23529.1; -.
CC      HSSP: P24941; 1BUH.
CC      MGD: MGI:97303; Nek1.
CC      InterPro: IPR000719; Euk_pkinase.
CC      InterPro: IPR002290; Ser_thr_pkinase.
CC      Pfam: PF00069; Pkinase; 1.
CC      SMART: SM00220; S_TKC; 1.
CC      PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC      PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC      PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC      TRANSFERASE: Serine/threonine-protein kinase; ATP-binding; Mitosis;
```

KW Nuclear protein; Phosphorylation; Cell cycle; Cell division;
 KW Tyrosine-protein kinase.
 FT DOMAIN 4 258 PROTEIN KINASE.
 FT NP_BIND 10 18 ATP (BY SIMILARITY).
 FT BINDING 33 33 ATP (BY SIMILARITY).
 FT ACT_SITE 128 128 BY SIMILARITY.
 SQ SEQUENCE 774 AA: 88427 MW: 7EDES881ACCC06FA CRC64:

Query Match 54.2% Score 3385; DB 1; Length 774;
 Best Local Similarity 85.5% Pred. No. 1, 2e-118;
 Matches 665; Conservative 45; Mismatches 64; Indels 4; Gaps 3;

QY 1 MEKYVRLQKIGSGFGKALIVKSTEDGRQYIKETINSRSSKREESREAVLANMKH 60
 DB 1 MEKYVRLQKIGSGFGKALIVKSTEDGRQYIKETINSRSSKREESREAVLANMKH 60
 QY 61 PNIVQYRESSEFNGSLVYMDYCEGDLFRKINAKGVLFQEDQILDMFVQICLAKHVN 120
 DB 61 PNIVQYRESSEFNGSLVYMDYCEGDLFRKINAKGVLFQEDQILDMFVQICLAKHVN 120
 QY 121 DKRIILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYLSPEICENKPY 180
 DB 121 DKRIILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYLSPEICENKPY 180
 QY 122 DKRIILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYLSPEICENKPY 180
 DB 122 DKRIILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYLSPEICENKPY 180
 QY 181 NKSQDMALGCVLYELCTLKHAFAGSMKNLYLKTISGSPRSLHYSDLSQLEK 240
 DB 181 NKSQDMALGCVLYELCTLKHAFAGSMKNLYLKTISGSPRSLHYSDLSQLEK 240
 QY 241 RNPDRPVSNTLEKFLAKRIEFLSPQIAEFELCTKTSKSGSPRIAPKAPASQNSI 300
 DB 241 RNPDRPVSNTLEKFLAKRIEFLSPQIAEFELCTKTSKSGSPRIAPKAPASQNSI 300
 QY 241 RNPDRPVSNTLEKFLAKRIEFLSPQIAEFELCTKTSKSGSPRIAPKAPASQNSI 300
 DB 241 RNPDRPVSNTLEKFLAKRIEFLSPQIAEFELCTKTSKSGSPRIAPKAPASQNSI 300
 QY 301 SVAPAKITPPAKYGIPLAYKKYGDKKLHKKRPLQKHKQAHOTPEKRVNTGEERRKISE 360
 DB 301 SVAPAKITPPAKYGIPLAYKKYGDKKLHKKRPLQKHKQAHOTPEKRVNTGEERRKISE 360
 QY 301 SEVPAPKIKRPAKAGVLYLTKYKGDKKLEKRRPRKQAHQIPYKKNMSGEERRKISE 360
 DB 301 SEVPAPKIKRPAKAGVLYLTKYKGDKKLEKRRPRKQAHQIPYKKNMSGEERRKISE 360
 QY 361 EAARKRLEFIEKERRKQDQITSLMAKQOMKROEKERLEINBARQGNRVLSAGSGE 420
 DB 361 EAARKRLEFIEKERRKQDQITSLMAKQOMKROEKERLEINBARQGNRVLSAGSGE 420
 QY 361 EAARKRLEFIEKERRKQDQITSLMAKQOMKROEKERLEINBARQGNRVLSAGSGE 420
 DB 361 EAARKRLEFIEKERRKQDQITSLMAKQOMKROEKERLEINBARQGNRVLSAGSGE 420
 QY 421 VNAFLGSGGTAPSSFGSGEYHYAIFPOMQOQAEENEAKKREITGRGLPERQNG 480
 DB 421 VNAFLGSGGTAPSSFGSGEYHYAIFPOMQOQAEENEAKKREITGRGLPERQNG 480
 QY 421 VNAFLGSGGTAPSSFGSGEYHYAIFPOMQOQAEENEAKKREITGRGLPERQNG 480
 DB 421 VNAFLGSGGTAPSSFGSGEYHYAIFPOMQOQAEENEAKKREITGRGLPERQNG 480
 QY 420 VNAFLGSGGTAPSSFGSGEYHYAIFPOMQOQAEENEAKKREITGRGLPERQNG 479
 DB 420 VNAFLGSGGTAPSSFGSGEYHYAIFPOMQOQAEENEAKKREITGRGLPERQNG 479
 QY 481 QLAVERAKOYEEFLQKREKAMONKARAGHYUATRIQRIQONENROQITAKIKGEK 540
 DB 481 QLAVERAKOYEEFLQKREKAMONKARAGHYUATRIQRIQONENROQITAKIKGEK 540
 QY 480 HLAVERAKOYEEFLQKREKAMONKARAGHYUATRIQRIQONENROQITAKIKGEK 539
 DB 480 HLAVERAKOYEEFLQKREKAMONKARAGHYUATRIQRIQONENROQITAKIKGEK 539
 QY 541 EANHSGOGESEADMRKKITLESKAMANARAVALKQDLERKREKAEVREKQWEEHVA 600
 DB 541 EANHSGOGESEADMRKKITLESKAMANARAVALKQDLERKREKAEVREKQWEEHVA 600
 QY 540 EADGTGQOAEETEDMRKKITLESKAMANARAVALKQDLERKREKAEVREKQWEEHVA 599
 DB 540 EADGTGQOAEETEDMRKKITLESKAMANARAVALKQDLERKREKAEVREKQWEEHVA 599
 QY 601 KGVKSDVSPRPGQHETGSGSPSKQOMRSVLSVTSALKEVGVSLSLDTRETSEEMKINN 660
 DB 601 KGVKSDVSPRPGQHETGSGSPSKQOMRSVLSVTSALKEVGVSLSLDTRETSEEMKINN 660
 QY 600 R-VKSDVSPRPGQHETGSGSPSKQOMRSVLSVTSALKEVGVSLSLDTRETSEEMKINN 656
 DB 600 R-VKSDVSPRPGQHETGSGSPSKQOMRSVLSVTSALKEVGVSLSLDTRETSEEMKINN 656
 QY 661 AISSKREILRLNENMLKAOEDEKQMONLSPFEINVEDAKHEHREKESVSDKKKEAG 720
 DB 661 AISSKREILRLNENMLKAOEDEKQMONLSPFEINVEDAKHEHREKESVSDKKKEAG 720
 QY 657 AISSKREILRLNENMLKAOEDEKQMONLSPFEINVEDAKHEHREKESVSDKKKEAG 716
 DB 657 AISSKREILRLNENMLKAOEDEKQMONLSPFEINVEDAKHEHREKESVSDKKKEAG 716
 QY 721 QLVIPRLDELTLST 778
 DB 721 QLVIPRLDELTLST 778
 QY 717 QLVIPRLDELTLST 774
 DB 717 QLVIPRLDELTLST 774

RESULT 2
 STR2_HUMAN
 ID STR2_HUMAN STANDARD: PRT: 841 AA.
 AC P51957;
 DT 01-OCT-1996 (rel. 34, Created)
 DT 01-OCT-1996 (rel. 34, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)

DE Serine/threonine protein kinase 2 (EC 2.7.1.37) (Serine/threonine-
 DE protein kinase NRK2).
 GN STR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_Taxid-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PREST:
 RX MEDLINE-9426838; PubMed-8208544;
 RA Lavedakou E.N., He M., Baptist E.W., Craven R.J., Cance W.G.,
 RA Welch F.L., Simons A., Naylor S.L., Leach R.J., Lewis T.B.,
 RA Bowcock A., Liu E.T.;
 RT Two novel human serine/threonine kinases with homologies to the cell
 RT cycle regulating Xenopus Mo15, and NIMA kinases: cloning and
 RT characterization of their expression pattern.";
 RL Oncogene 9:1977-1988(1994).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADULT HEART, FOLLOWED BY
 CC PANCREAS, SKELETAL MUSCLE, BRAIN, LIVER, KIDNEY, LUNG AND
 CC PLACENTA. PRESENT IN MOST PRIMARY CARCINOMAS.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIMA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 CC EMBL: L20321; AAA36558.1;
 CC HSSP: P11362; 1FGK.
 CC MIM: 601959;
 CC DR InterPro: IPR000719; Euk_pkinase.
 CC DR InterPro: IPR002290; Ser_thr_pkinase.
 CC DR Pfam: PF00069; pkinase; 1.
 CC DR SMART: SM00220; S_TKc; 1.
 CC DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 CC DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
 CC DR PROSITE: PS00111; PROTEIN KINASE DOM; 1.
 CC KW Transferase: Serine/threonine-protein kinase; ATP-binding; Mitosis;
 KW Nuclear protein; Phosphorylation.
 FT DOMAIN 6 261 PROTEIN KINASE.
 FT NP_BIND 12 20 ATP (BY SIMILARITY).
 FT BINDING 35 35 ATP (BY SIMILARITY).
 FT ACT_SITE 131 131 BY SIMILARITY.
 FT MOD_RES 165 165 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT SPOUNCE 841 AA: 94571 MW: 0DD31920DDE7E58 CRC64;

Query Match 13.6% Score 847.5; DB 1; Length 841;
 Best Local Similarity 28.1% Pred. No. 5, 9e-25;
 Matches 263; Conservative 160; Mismatches 316; Indels 197; Gaps 34;

QY 4 YVRLQKIGSGFGKALIVKSTEDGRQYIKETINSRSSKREESREAVLANMKHPNI 63
 DB 4 YVRLQKIGSGFGKALIVKSTEDGRQYIKETINSRSSKREESREAVLANMKHPNI 63
 QY 6 YCYLRVYVNGSGYGVYLVKRRDQKQYIKETINSRSSKREESREAVLANMKHPNI 65
 DB 6 YCYLRVYVNGSGYGVYLVKRRDQKQYIKETINSRSSKREESREAVLANMKHPNI 65
 QY 64 VQYRESFE-ENGLSVYMDYCEGDLFRKINAKGVLFQEDQILDMFVQICLAKHVN 122
 DB 64 VQYRESFE-ENGLSVYMDYCEGDLFRKINAKGVLFQEDQILDMFVQICLAKHVN 122
 QY 66 VYTESMEGGDGLYVMDYCEGDLFRKINAKGVLFQEDQILDMFVQICLAKHVN 125
 DB 66 VYTESMEGGDGLYVMDYCEGDLFRKINAKGVLFQEDQILDMFVQICLAKHVN 125
 QY 123 KILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYLSPEICENKPY 182
 DB 123 KILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYLSPEICENKPY 182
 QY 126 HILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYLSPEICENKPY 185
 DB 126 HILHROIKSONIFLTGDTGVQGLDFGIARVLSVTELARTCIGTPYLSPEICENKPY 185
 QY 183 KSDVWALGCVLYELCTLKHAFAGSMKNLYLKTISGSPRSLHYSDLSQLEK 242
 DB 183 KSDVWALGCVLYELCTLKHAFAGSMKNLYLKTISGSPRSLHYSDLSQLEK 242
 QY 186 KSDVWALGCVLYELCTLKHAFAGSMKNLYLKTISGSPRSLHYSDLSQLEK 245
 DB 186 KSDVWALGCVLYELCTLKHAFAGSMKNLYLKTISGSPRSLHYSDLSQLEK 245

QY	243	PDRPVSNIISKEKFIKRLIEKLSPQILAEIEFCIKTFKSGSOPITAKRPAKSGNSISV	302
QY	243		302
Db	246	PERPESVSIILROYIKROIISFLE----	281
QY	303	MPAQKITPKPAKYGIPLAYKKYIDDKLHKKPL-----	353
Db	282	-----KNDSS--OSKPFATVSGEASNHEVTHPQ----	310
QY	354	ERRKIISEPAARKRL--EFTEKEKKOKDOI--SLMAQDMQRQ-----	402
Db	311	---LSESSQYIIMGECKCISOEKRPAASGLPASKATKATKODISNTLTLATISSV	366
QY	403	RAREQWRNVLSAGSGEGEYKAPFLGSGGITAPSSSSKQYENHATFDMOQOARND	462
Db	367	-----IDILPAKRDSVSGFQVOEN--QPRYIDANIELGIGICSI--SVEEEMLDQMT	415
QY	463	AK-----YKKEIYGRLEPRQNG--QLAVEAKQVEEFLQKREMANONKARAEHGM	511
Db	416	KSSAQENILPMSSSIDYTGKKNPPVAPQPILEKQKPDGL-----ALSPKIECSGTI	470
QY	512	VYIARLROILRNFQNRQIQIAKRLGKRRKANISEQESSEFADMRKKEI-SLKAAHVA	570
Db	471	LAISNL--RLIGSSDSPASASVAVGITGVCHIAQOVAVGECITIKQRIHPILOPINS	527
QY	571	RAAVIKLEOLERRKEKEYEKKWEEMHVAKVGSADVPP--LGQHTGGSPEKQOORS	628
Db	528	SEPSLSNQRKRRKQTEHR--GEKROVRDILFAQESPPRFLPSHPITYEK-----	576
QY	629	VIVSYSLAKRVG----VDSILDTQRTRESEMKQTNNASIKRELIRLNLNKAQDEK	683
Db	577	-VDSTSTOKAEORNRVYVGVSSSR--SESMSSKDRPLSARE--RRLKQ---SDEMS	629
QY	684	GMQMLSTPEFINVHEPAKHEKEK-----SVSSDRKKWEAGQIVIPIDETLD	732
Db	630	SSGSYVAKALSVAGGKFOEEDVDPALRRLSDSCVSTQERKQIHC-----LSEDESS	684
QY	733	TSFTTTRHRTVGEVINKLGPNGSPFRAMGSEFDSVYLKIGEAELQTEBL-----	784
Db	685	TSSTDKSDGDYGE-----GKGQTNEL-----NALVQLMTQTLKDSKESCE	725
QY	785	NTTIRSEIS--PEGEKYKPLTIGEKKVCISHEIN--PSAIVDSPVETKSPESEA--	836
Db	726	DVPANPVSSEPKLRHRTKOTLLHKKVAEAEELHFEFLPSALM--PGEKIRRLVEYLR	789
QY	837	-----SPQSKLKEGNEEEDDLETEI--LOEBSG	864
Db	784	TDVIRGLAGVQLDEQVYDILDEEDREDREVRLEHMG	819
RESULT	3		
TK	NEK3 MOUSE	STANDARD:	PRG: 511 AA.
AC	16-OCT-2001: 0920X9:		
DT	16-OCT-2001 (Rel. 40, last sequence update)		
DT	16-OCT-2001 (Rel. 40, last annotation update)		
DE	Serine/threonine-protein kinase NEK3 (EC 2.7.1.1.) (NIMA-related protein kinase 3).		
GN	NEK3.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
ON	NCBI_Taxid=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE-99240743: PubMed-10224116;		
RA	Tanaka K., Nigg E.A.;		
RT	Cloning and characterization of the murine Nek3 protein kinase, a		
RT	novel member of the NIMA family of putative cell cycle regulators.;		
RT	J. Biol. Chem. 274:13491-13497(1999).		
RN	[2]		
RP	SEQUENCE FROM N.A.		

Query Match	13.3%	Score 833	DB 1	Length 511
Best Local Similarity	37.4%	Pred No 1.2e-24	Mismatches 174	Indels 88
Matches 203	Conservative 78	Identical 174	Indels 88	Gaps 16
DB	1	MEKRYRLQIGSGSPGRKALYKSTFSDROVYIKINISMSSKERESESRREAVLANMKH	60	
QY	1	MEKRYRLQIGSGSPGRKALYKSTFSDROVYIKINISMSSKERESESRREAVLANMKH	60	
DB	1	MDNYVLAFLVIGSGRRLLVLOESSNQTPAKMEI---PLKSDPQTSRKEAVLLAKMKH	57	
QY	61	PNIVQYSESEFNGSLVYMDYCGDGLFRINAOKGVLFQEQDILDMFVQICLALKH	120	
DB	58	PNIVAFSEFEAEGLVYIMVYCGDGLMORIKOOKGNFLPEPTILNMFIOICLGVNHH	117	
QY	121	DKRLIHRODKSONIFLTFMDQTVQDGDGIRAVNSTVTELARTCIGTPYVLSPEICENPY	180	
DB	118	KRVLEHRODKSNVLEITNGKVKVLLKGDGSRALLSSPFAFCYVGGPYVPEIWEENPY	177	
QY	181	NNKSDIWAIGCVLYELCTLKHAFAGSAGMKNLVLIKISGSPPVSLHAYSDRLSVQLFK	240	
DB	178	NNKSDIWAIGCVLYELCTLKHAFAGSAGMKNLVLIKISGSPPVSLHAYSDRLSVQLFK	237	
QY	241	RNRDRPVSNSLIEGFIATRIEKLFSQLIAE-----EFLCKTSPKFSQPIPAKR	293	
DB	238	RNSHRPVSNTLLKCRSLIAPVLPKCLPQIIIREYGEQIDELIKISTPKMKKQDSNRGR	297	
QY	294	ASQNSNISVMPQKI-----TKDA-----AKGIPLAKKYQYKLIH--EK	332	
DB	298	ALPEANSAAMQEEERKCRKSTHLESTGTTPAGNALORARGN-----ESGNGDEGSHIT	353	
QY	333	KPLQKRGAAHQT--PEKRYN-----TEEEKRKS-----DEAARKRLTEIEKE	374	
DB	354	SPASPHPHPRWHRHGSPSNVEALEKASITSTSFADDDRGSGVYIKEDNARQVY-----	407	

OY 375 KQKQDQIISLMKAQOMKROEKERLERINRAREQGNVLSAGSGEYKAPLGGCTIAR 434
 DB 408 REPPEALLSMKQADLSQAFQF-----YTIYRGAEGEFLKGLP--SEDFASD 452
 OY 435 SFSFSSRQGY----EHYHAFEDMOOQRAEDNE-AKKKREI-----YGRGEPERQKGLAV 484
 DB 453 SYDGDLDSDVWLPDEREPRLDEEDTDDEEDNENPDWVSELKKHVGYGDC-----PGGLLG 508
 OY 485 ERA 487
 DB 509 ERA 511
 RESULT 4
 NRK3_HUMAN STANDARD: PRT; 459 AA.
 AC P51366;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase NEK3 (EC 2.7.1.1.) (NIMA-related
 protein kinase 3) (HSPK 36) (Fragment).
 GN NEK3
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE=94368699; PubMed=7522034;
 RA Schultz S.J., Fry A.M., Suetterlin C., Ried T., Nigg E.A.;
 RT "Cell cycle-dependent expression of Nek3, a novel human protein
 kinase related to the NIMA mitotic regulator of Aspergillus
 nidulans";
 RT Cell Growth Differ. 5:625-635(1994).
 (2)
 RN Cell Growth Differ. 5:625-635(1994).
 RP MEDLINE=94100173; PubMed=8274451;
 RA Schultz S.J., Nigg E.A.;
 RT "Identification of 21 novel human protein kinases, including 3
 members of a family related to the cell cycle regulator nima of
 Aspergillus nidulans";
 RT Cell Growth Differ. 4:821-830(1993).
 RL -1- FUNCTION: KINASE THAT MAY PLAY A ROLE IN MITOTIC REGULATION.
 CC -1- DISEASE: MIGHT BE CANDIDATE FOR USHER SYNDROME, BECAUSE OF ITS
 CC CHROMOSOMAL LOCATION.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIMA SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 DB EMBL: Z29067; CA82310.1;
 DB EMBL: Z2534; CA80921.1;
 DB HSP: P24941; ICRP.
 DB MIM: 604044; ICRP.
 DR InterPro: IPR000719; Euk-kinase.
 DR InterPro: IPR002230; Ser-THR-kinase.
 DR Pfam: PF00069; Kinase; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KM Transferase: Serine/threonine-protein kinase; ATP-binding; Mitosis;
 KW Nuclear protein; Phosphorylation; Cell cycle; Cell division.
 FT NON_TER 1
 FT DOMAIN 1
 FT ACT_SITE 80 210
 FT BY SIMILARITY.

FT MOD_RES 118 118 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CONFLICT 30 31 IV -> LY (IN REF. 2).
 FT CONFLICT 140 142 SLG -> PSV (IN REF. 2).
 SQ SEQUENCE 459 AA; 52300 MW; DB506EC30EAB49 CRC64;
 Query Match 12.7%; Score 790.5; DB 1; Length 459;
 Best local Similarity 42.9%; Pred. No. 4e-23;
 Matches 156; Conservative 58; Mismatches 87; Indels 63; Gaps 5;
 OY 49 REVAVLANKMHPNIVORYRESPEENGSLTYVMDCYCEGGLDFKRINAKGVLFQEDQIIDM 108
 DB 1 KREAVLSAKMKHPNIVAFKESFEAGHLTYVMERYCGGGLDMKIKQCKLPEDMIINW 60
 OY 109 FVOICLALKVHDKRIIHRDIKSONIFLTKDGTVOGLDGIARVNSTVELARTCIGTPY 168
 DB 61 FTOMCLGVNHIKKRVLRDIKSNIFLTONCKVAKLGDGFSARLLSNPAFACTVGTPT 120
 OY 169 YLSEICENKRYNKKSDIMALGCVLYELCTLKHAFAGSMKNLVYIKTISGFPVSLHS 228
 DB 121 YVPEIENLPIYKNSDIMSGLILECTLKHFPQANSKMLIKVQCGCISPLSHYS 180
 OY 229 YDLRSIYSLFKRPPDRPVSNSILEKFTAKRIEKLSPQILAEFCLTKTFKFSQPI 288
 DB 181 YELQFLVKQKFRNPSHRFSATTLNLSRGIVARVQCKLPETLME----- 225
 OY 289 PAKRPASGNSISVMPAKITKPAKYGIPLVAKYGGKKLHKKPLOKHQAQTPK 348
 DB 226 -----YGEVLEIK-----NSKNTPRKK 245
 OY 349 VNTGEERRKISEEAKRRRLFEFEKEKKQKQDQIISLMKAQOMKROEKERLERINRAREQ 408
 DB 246 TNSRIRIALGNASIVQ-----EEQDKKSGHTDELSINE--NLYESALNRVNR-EK 297
 OY 409 WRNV 412
 DB 298 NKS 301
 RESULT 5
 NRKA_TRYBB STANDARD: PRT; 431 AA.
 ID AC 008942;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative serine/threonine-protein kinase A (EC 2.7.1.37).
 GN NRKA.
 OS Trypanosoma brucei brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_Taxid=5702;
 (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ETRO 164; ISOLATE ISTARI;
 RX MEDLINE=93295429; PubMed=8515773;
 RA Vale M.J., Jr., Parsons M.;
 RT "A trypanosoma brucei gene family encoding protein kinases with
 RT catalytic domains structurally related to Nek1 and NIMA";
 RT Mol. Cell. Biochem. Parasitol. 59:111-122(1993).
 RL -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 DB EMBL: L03778; AB59252.1;
 DB HSP: P24941; IAO1.

Query Match	9.1%;	Score 565.5;	DB 1;	Length 431;
Best Local Similarity	40.5%;	Pred. No. 7.6e-15;		
Matches 109; Conservative	63;	Mismatches 92;	Indels 5;	Gaps 2

RESULT	6	
NEK2_HUMAN		
ID	NEK2_HUMAN	STANDARD;
		PRT;
		445 AA

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, last sequence update)
DT 16-OCT-1996 (Rel. 40, last annotation update)
DE Setine/threonine-protein kinase NRK2 (EC 2.7.1.-) (NIMA-related
DE protein kinase 2) (Nima-like protein kinase 1) (HSPK 21).
GN NRK2 OR NLU1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=T-cell, placenta, and Mesopharynx;
RC MEDLINE=94368699; PubMed=1522031; C., Ried T., Nigg E.A.;
RA Schmitt S.J., Fry A.M., Swetlow J.C., expression of NrK2, a novel human protein
RT "cell cycle-dependent expression of NrK2, a novel human protein
RT kinase related to the NIMA mitotic regulator of Aspergillus
RT nidulans.";
RL cell growth differ. 5:625-635(1994).
RN [2]
RP SEQUENCE FROM N.A.
RP Lu K.P., Hunter T.;

```

RT  "Molecular cloning and expression of NIK1, a human NIMA-like kinase.";
RL  Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
RN  [31]
RP  SEQUENCE OF 83-203 FROM N.A.
RX  MEDLINE=94100173; PubMed=8274451;
RA  Schultz S.J., Nigg E.A.;
RT  "Identification of 21 novel human protein kinases, including 3 members
RT  of a family related to the cell cycle regulator NIMA of Aspergillus
RT  nidulans.";
RL  Cell growth Differ. 4:821-830(1993).
CC  -1- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN MITOTIC REGULATION,
CC  MAY HAVE A ROLE AT THE G2-M TRANSITION. MAY ALSO PLAY A ROLE IN
CC  MEIOSIS.
CC  -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC  -1- DEVELOPMENTAL STAGE: ACCUMULATES THROUGHOUT S PHASE AND SHOWS
CC  MAXIMAL LEVELS IN LATE G2. THIS EXPRESSION PATTERN IS HIGHLY
CC  REMINISCENT OF THAT OF A AND B CYCLINS.
CC  -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC  NIMA SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; Z29066; CAAB2309.1; -
CC  EMBL; 011050; AA819558.1; -
CC  EMBL; Z25425; CAAB0912.1; -
CC  DR  HSSP; P00518; IPRK.
CC  MI; 604043; -
CC  DR  InterPro; IPR000719; Euk_Pkinase.
CC  DR  InterPro; IPR002290; Ser_thr_Pkinase.
CC  DR  Pfam; PF00069; pkinase; 1.
CC  SMART; SM00220; S_TKC; 1.
CC  DR  PROSITE; PS00107; PROTEIN_KINASE_ATP, FALSE_NEG.
CC  DR  PROSITE; PS00108; PROTEIN_KINASE_SF; 1.
CC  DR  PROSITE; PS00114; PROTEIN_KINASE_DOM; 1.
CC  DR  Transferr; Serine/threonine-protein kinase: ATP-binding; Mitosis;
CC  Nucleosome phosphorylation; Cell cycle; Cell division; Meiosis.
CC  KM  Nucleosome protein; phosphorylation; Cell cycle; Cell division; Meiosis.
CC  KM  DOMAIN 8 271
CC  KM  NP_BIND 14 22
CC  KM  BINDING 37 37
CC  KM  ACT_SITE 141 141
CC  KM  MOD_RES 179 179
CC  KM  CONFLICT 85 85
CC  FT  IV -> LY (IN REF. 3).
CC  SEQUENCE 445 AA; 51763 MW; D33A3778ABBD9E CRC64;

```

DB 244 LNEITITRLMLNLDKVRPSVEEILENPLIA-----DLVADE-ORRLNERGRG--L 290
 QY 291 KRPAQSGNSISVMPAOKITKPAANYGIPLAYKKYGDKKLHEKRPLOKHQAOTPEKRYN 350
 DB 291 GEPEKSDSSPVLSELT-----KEIQDEREALAKAREERLE 328
 QY 351 TGPE-----RRKISEE-AARKRRRL--EFIEKEK-----QKDOIISLMKAEQKKR----- 392
 DB 329 QKQDELQVARELAEDEKLARAENLKNYSILKEREFTSLANPELNLPPSVYKKKVFSG 368
 QY 393 QEKERLERINRAREQ 407
 DB 389 ESKENIMKSENSESO 403

RESULT 7
 ID NEK2_MOUSE STANDARD: PRT: 443 AA.
 AC 035942; 035959;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Serine/threonine-protein kinase Nek2 (EC 2.7.1.-) (Nima-related protein kinase 2).
 GN NEK2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;

RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
 RC STRAIN=SWISS WEBSTER; TISSUE=Testis;
 RA MEDLINE=9/330684; PubMed=9187143;
 RA Rhee K., Wolgemuth D.J.;
 RT "The NIMA-related kinase 2, Nek2, is expressed in specific stages of
 RT the meiotic cell cycle and associates with meiotic chromosomes";
 RT Development 124:2167-2177(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98243037; PubMed=9583679;
 RA Arama E., Yanai A., Kilfin G., Motio B.;
 RT "Murine NIMA-related kinases are expressed in patterns suggesting
 RT distinct functions in gametogenesis and a role in the nervous
 RT system";
 RT Oncogene 16:1813-1823(1998).
 RN [3]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=Brain;
 RA MEDLINE=98096235; PubMed=9434622;
 RA Tanaka K., Parvlien M., Nigg E.A.;
 RT "The in vivo expression pattern of mouse Nek2, a NIMA-related kinase,
 RT indicates a role in both mitosis and meiosis";
 RT Exp. Cell Res. 237:264-274(1997).
 RL -!- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN MITOTIC REGULATION.
 CC MAY HAVE A ROLE AT THE G2-M TRANSITION. MAY ALSO PLAY A ROLE IN
 CC MEIOSIS.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: MOST ABUNDANTLY EXPRESSED IN TESTIS. LOW
 CC LEVELS FOUND IN MID-GESTATION EMBRYO, OVARY, PLACENTA, INTESTINE,
 CC THYMUS AND SKIN. WITHIN THE TESTIS, EXPRESSION RESTRICTED TO GERM
 CC CELLS WITH HIGHEST LEVELS DETECTED IN SPERMATOCTES AT PACHYTEME
 CC AND DIPTYOTENE STAGES. ALSO EXPRESSED IN MEIOTIC PACHYTEME OOCYTES.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIMA SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U95610; AAB67973.1; -
 DR EMBL: AF013166; AAC35393.1; -
 DR EMBL: AF007247; AAB70470.1; -
 DR HSSP: P24941; 1CRP.
 DR MGD: MG1:109359; Nek2.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR SMART: SM00220; S-TRK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Transferase: Serine/threonine-protein kinase; ATP-binding; Mitosis;
 KW Nuclear protein; Phosphorylation; Cell cycle; Cell division; Meiosis;
 FT DOMAIN 8 271
 FT NP_BIND 14 22 ATP (BY SIMILARITY).
 FT BINDING 37 37 ATP (BY SIMILARITY).
 FT ACT_SITE 141 141 BY SIMILARITY.
 FT MOD_RES 179 179 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 SQ SEQUENCE 443 AA: 51307 MW: DE09565C378307E1 CRC64;

Query Match 8.7%; Score 544.5; DB 1; Length 443;
 Best local similarity 27.7%; Pred. No. 4.6e-14;
 Matches 152; Conservative 85; Mismatches 182; Indels 129; Gaps 14;

QY 1 MEKYVRLQKIGSGFGKALVYKSTEDGROYIKRINISMSKERESESRVAVLANMKH 60
 DB 5 VEDIEVLSHIGSTGRCKIRKSDGKILVKELDGSMTEVEKQMLSEVNLRELKH 64
 QY 61 PNIVQYRESF--EENGSLIYMDYCEGDLFRKINAQV---LROEDQILDMPVOITL 114
 DB 65 PNIVRYDYIDITRTVTTLYIWEYCEGGDLASVIS--KGRKQROYLEEVEVLVMTQLT 122
 QY 115 ALKHVDRK-----ILHRDKSQNIFLTKDGTVOAGDGIARVNLSTVELARTICFPY 169
 DB 123 ALKECHRRRDGSHVTLHRDLKPAVNFILDSKHNVLGDFGLARILNHDTSPAKTFVOTPY 182
 QY 170 LSPFICENRPYNNKSDIYALGCVLELCTLKHAEASMKNLVLIIGSFPPVSILHYSY 229
 DB 183 MSPEQSGCYSTYNEKSPDYSWLMACLELCAIMPPTAFARNOKELAGKIREGFRRIPIRYSD 242
 QY 230 DLRSVLSOLFKNPNDRPSVNSILEKGFIAKRIKFLSPOLIAEEFCIKFSGSQIP 289
 DB 243 GLNDLITRMFLKDYHRPSVEEILESPLIA-----DMVAEE----- 278
 QY 290 AKRPASGNSISVMPAOKITKPAANYGIPLAYKKYGDKKLHEKRPLOKHQAOTPEKRY 349
 DB 279 -----ORRLNERGRSGEP-----SKLPDSSPVLSEU--KLKESQL 313
 QY 350 NTGEERRKISEBARRRRLFEIEKEKQKQDI--SLMKAEQMKROEKRELERINRARE 406
 DB 314 QDEBQALRAREDIILRQKEREICIRERLEDEKLAAEESIMNYSILKEHR----- 362
 QY 407 OGRNVLASGSGEYVAPLFGSGTIAPSSFSFGQEHYALFDQMOQOARADNAKKK 466
 DB 363 -----LLCLAGGEPLDLP-----SSAKKKVYHFG-----ESKENTA--- 394
 QY 467 RELYGGLEPERQQLAVERAKQVEFLORREREMOKMARAEGHYMLARLRIORLON 526
 DB 395 -----RSNSESYLAK-----SKRDKRKRLLHAOLRQALADLE 429
 QY 527 EROQITAK 534
 DB 430 KNTOLKSR 437
 RESULT 8
 ID NIMA_EMEI STANDARD: PRT: 699 AA.
 AC P11837;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 GN G2-specific protein kinase NIMA (EC 2.7.1.1.) (Never in mitosis).
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eukaryotes; Trichocomaceae; Emericella.
 OX NCBI_TaxID=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=68194523; PubMed=3359487;
 RA Osmari S.A., Fu R.T., and Morris N.R.;
 RT "Mitotic induction and maintenance by overexpression of a G2-specific
 RT gene that encodes a potential protein kinase.";
 RT Cell 53:237-244(1988).
 RL [2]
 RP CHARACTERIZATION
 RX MEDLINE=95196762; PubMed=7889945;
 RA "R.K.T., Osmari S.A.";
 RT "Isolation of a functional homolog of the cell cycle regulated NIMA protein kinase
 RT of Aspergillus nidulans is required for mitotic exit.";
 RT PMO J. 14:995-1003(1995).
 RL [3]
 RN MUTAGENESIS.
 RX MEDLINE=95355415; PubMed=7629122;
 RA Pu R.T., Xu G., Wu L., Vierula J., O'Donnell K., Ye X.S.,
 RT Osmari S.A.;
 RT "Isolation of a functional homolog of the cell cycle-specific NIMA
 RT protein kinase of Aspergillus nidulans and functional analysis of
 RT conserved residues.";
 RT J. Biol. Chem. 270:18110-18116(1995).
 RL [4]
 RP FUNCTION: PROTEIN KINASE THAT PLAYS AN IMPORTANT ROLE IN MITOTIC
 RT REGULATION. SEEMS TO BE PHOSPHORYLATED AND THEREBY ACTIVATED BY
 RT CDC/CYCLIN B DURING MITOTIC ACTIVATION. IT IS ALSO REQUIRED FOR
 RT SPMIDE FORMATION AND FOR NUCLEAR ENVELOPE BREAKDOWN.
 CC [1] SUBCELLULAR LOCATION: Nuclear (Probable).
 CC [1] DEVELOPMENTAL STAGE: ACCUMULATES WHEN CELLS ARE ARRESTED IN G2
 CC AND IS DEGRADED AS CELLS TRAVERSE MITOSIS.
 CC [1] PTM: RECESSIVE MUTATIONS OF NIMA CAUSE A SPECIFIC CELL CYCLE BLOCK
 CC IN G2 AT RESTRICTIVE TEMPERATURE DUE TO LACK OF PHOSPHORYLATION
 CC THAT NORMALLY ACTIVATES THE G2/KINASE.
 CC [1] SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIMA SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M20249; AAA33316.1;
 DR PIR: A43734; A43734.
 DR HSSP: P05132; OCT19; Euk_pkinase.
 DR InterPro: IPR001220; Ser_Thr_pkinase.
 DR Pfam: PF00063; PK_kinase.1.
 DR SMART: SM00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_SF.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 DR KinaseBase: Serine/threonine-protein kinase; ATP-binding; Mitosis;
 DR KinaseBase: Serine/threonine-protein kinase; Cell cycle; Cell division.
 KM Domain: 11 295
 RN NP_BIND 17 25
 FT BINDING 40 40 ATP (BY SIMILARITY).
 FT ACT_SITE 166 166 ATP (BY SIMILARITY).
 FT MOD_RES 199 199 PHOSPHORYLATION (AUTO-).
 FT MUTAGEN 38 38 C->S/A: NO LOSS OF ACTIVITY.
 FT MUTAGEN 41 41 E->G: IN NIMA7; BLOCK IN G2.
 FT MUTAGEN 91 91 Y->N: IN NIMA5; BLOCK IN G2.
 FT MUTAGEN 199 199 T->A: GREAT DECREASE IN ACTIVITY.
 FT MUTAGEN 304 304 L->P: IN NIMA1; BLOCK IN G2.

SQ SEQUENCE 699 AA; 78912 MW; 50FDC86E249ABC9 CRC64;
 Query Match 8.5%; Score 530.5; DB: 1; Length 699;
 Best Local Similarity 27.8%; Pred: No 2.4e-13;
 Matches 168; Conservative 92; Mismatches 206; Indels 139; Gaps 18;
 OY 2 EKYVRLQKTEGSGFKALVYKSTEDGROVIVKINISMSKREKRSREVAVLANKHP 61
 DB 9 DKVELEKIGSGSGRIIRKVKRSDGFIICREKINIKSTERRQGLAEFNILSRHP 68
 OY 62 NIVQV--RESEENGSLIXIMDCDEGDD- - - - -FKRINAGKVLPEQDILMFQVQIC 113
 DB 69 NIVATYHREHILASQDILYIMETCGGGDLSMTVKNLKRNYA- - - - -DEDFVMTLSQLV 124
 OY 114 IALKVH- - - - -DKRIHRDKSQNIFLFRDGVQLGD 146
 DB 125 TALYRCHIGTDPAYGNSLGPAPKPSGLKQAOQMTILRDLKPEINIFLGSNDTVKLGD 184
 OY 147 FGIARVLSVTEIARCTGIPPTYSPEICENKPYNNKSDIMALGCVLELTCLKHAFAG 206
 DB 185 FGLSKLHSH-DEASTYGTPTFYSPEICAEKRYTLRSDIMAVAGCIMELCORPPFNAR 243
 OY 207 SAKNIVATKISGPPPVSLAYSVDLSVQLFKRNPDRDPSVNSILKGFATKRIEKL 266
 DB 244 THIQVQKTRCKFAFLPDEFSSEKKNVIVASCLRVNPDHDPATLILNPVI- - - - - 295
 OY 267 SPQLAEFECLTFKPSGQPTPAKRPASGNSISVPAQKIKPKAKYGIPLAVAKYQD 326
 DB 296 - - - - -RLMRREVLNLS- - - - -RAARKREVTQKAKDVQAFK- - - - - 331
 OY 327 KRIHEKPLQKROKHOPPEKRVNTGEERRKISPEAKRRRLFEIKERKOKDQIISLAK 386
 DB 332 - - - - -LEKKEQKRS- - - - -ELNSIRREWEVKARLIDQVQNEIDKLKRF 375
 OY 387 AEQMKR- - - - -DEKRELEINRAREQGRNVLISAGSGEYKAPFLSGGTIAPSSFGQYE 444
 DB 376 CEVDRAQAEVEKORANRYREDA- - - - -SLRSSGHSQ- - - - -SSNSDSDDPSST- - - 425
 OY 445 HYAIFQMOQCAEDNEAKRWRETYGRLDPRQKGLAVAKQVEFLQRRKEMOK 504
 DB 426 - - - - -ISQLSLESPYTKAKK- - - - -LPKKE-SRPTFKSTVY- - - - -DSPMDI 463
 OY 505 ARAEGHWVYLARLQRIQNFNERQIKAKLGEKKKANSGEQSGSEEDMRKKEI 564
 DB 464 QMAEPPSISTASL- - - - -SLSPRTSATYS-GNITRADEGERRRRKPEFT 506
 OY 565 KAHAN 569
 DB 507 LAYSD 511
 RESULT 9
 NRKB_TRYBB STANDARD; PRT; 431 AA.
 ID NRKB_TRYBB
 AC 003428;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative serine/threonine-protein kinase B (EC 2.7.1.37).
 GN NRKB
 OS Trypanosoma brucei brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5702;
 RN SEQUENCE FROM N.A.
 RP STRAIN-ISOLATE TREU66;
 RX MEDLINE=93295429; PubMed=8515773;
 RC Gale M.J., Jr., Parsons M.;
 RA "A trypanosoma brucei gene family encoding protein kinases with
 RT catalytic domains structurally related to Nekl and NIMA.";
 RT Mol. Biochem. Parasitol. 59:111-122(1993).
 CC [1] CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.

```

CC CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: L03777; AA859253.1;
CC CC HSSP: P24941; IAQ1.
CC CC InterPro: IPR000719; Euk_pkinase.
CC CC InterPro: IPR001849; PH.
CC CC Pfam: PF00169; Ser_thr_pkinase.
CC CC Pfam: PF00231; Kinase.1.
CC CC SMART: SM00220; SH.1.
CC CC PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
CC CC PROSITE: PS00108; PROTEIN_KINASE_ST.1.
CC CC PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
CC CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
CC CC TRANSFERASE; Serine/threonine-protein kinase; ATP-binding.
CC CC DOMAIN 20 279
CC CC NP_BIND 331 429
CC CC BINDING 26 34
CC CC ACT_SITE 147 147
CC CC ACT_SITE 147 147
CC CC BY SIMILARITY.
CC CC SEQUENCE 431 AA; 48172 MW; F6B4B3E9E9E9F74 CRC64;

Query Match
Best Local Similarity 8.5%; Score 528.5; DB 1; Length 431;
Matches 103; Conservative 62; Mismatches 99; Indels 5; Gaps 2;

QY 3 KYVRLKIGESFGKALIVKSTEDGRQYIKENIRMSKEREESREAVANMKHPN 62
DB 19 KYLNGKIVIGSGYGEVVAERVDGSLCAKAWDLKSKMRKRYAQSIEIKYPTNCNHPN 78
QY 63 IVQYRESFENGSLYIVMDYCEGGDLFKRI---NMQKGLNEDDILDMFOICLAKHV 119
DB 79 IIRYIEDHEENDRLIYMEFADSGNLDEQIKPMGTGDARLYRHEHALFLFQLDCLALDYI 138
QY 120 HDRKILHRDIKSONIFLTKDGTVOGLDFGIARVNSTVE--IARCTIGTPYLSPECEN 177
DB 139 HSHKMLHRDIKSNVLTFTGLVLCDFGSHQYEDIVSGVASFQCTPYTLAPELMNN 198
QY 178 KPYNNKSDIWAQCVLELCTLKHAFAEGSMKNLVKTIISGSPVSLHYTDLRSLYSQ 237
DB 199 LRYNKRADVWSLGLVLLYEIMGMKRPFSASNLKGLMSKVLATYAPLDPSESEFKRYVDC 258
QY 238 LFRNRPDRPSVNSILEKGFIAKRIEKL 266
DB 259 ILVADNDRPSVNEFNQIYINKGLFLV 287

RESULT 10
GIN4_YEAST
GIN4_YEAST STANDARD: PRT; 1142 AA.
AC 012263:
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAR-2000 (Rel. 39, Last annotation update)
GN Serine/threonine-protein kinase GIN4 (EC 2.7.1.1.-).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_taxonomy:4932;
RN NCBI_taxonomy:4932;
RP SEQUENCE FROM N.A.

RX MEDLINE:99030835; PubMed:9813093;
RT Longtine M.S., Fares H., Pringle J.R.;
RT "Role of the yeast Gln4p protein kinase in septin assembly and the
RT relationships between septin assembly and septin function.";
RT J. Cell Biol. 143:719-736(1998).
(2)
RN RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Benito A.,
RA Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA Hunicke-Smith S., Hyman R., Komp C., Laskari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Oeffner P., Oh C., Petel F.X.,
RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA Wiant A., Yelton M., Botstein D., Davis R.W.;
RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MAY PLAY A ROLE IN SEPTIN ASSEMBLY.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CC NIM1 SUBFAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: U33140; AA75513.1;
CC CC HSSP: U33057; AA64949.1;
CC CC HSSP: P36897; ITRH.
CC CC SGD: S0002915; GIM4.
CC CC InterPro: IPR000719; Euk_pkinase.
CC CC InterPro: IPR002290; Ser_thr_pkinase.
CC CC Pfam: PF00069; pkinase.1.
CC CC SMART: SM00220; SH.1.
CC CC PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
CC CC PROSITE: PS00108; PROTEIN_KINASE_ST.1.
CC CC PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
CC CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
CC CC TRANSFERASE; Serine/threonine-protein kinase; ATP-binding.
CC CC DOMAIN 19 289
CC CC NP_BIND 25 33
CC CC BINDING 48 48
CC CC ACT_SITE 156 156
CC CC ACT_SITE 156 156
CC CC BY SIMILARITY.
CC CC SEQUENCE 1142 AA; 129857 MW; EC16FF4B49DD811 CRC64;

Query Match
Best Local Similarity 8.0%; Score 502.5; DB 1; Length 1142;
Matches 252; Conservative 215; Mismatches 475; Indels 331; Gaps 42;

QY 8 QKIGESFGKALIVKSTEDGRQYIKENIRMSKEREESREAVANMKHPN 62
DB 23 ETLGIGSTGQVQLARNGSTGOEAIVKSKAVNTGNVSTIGSTTDPALPYGIENR 82
QY 53 AVLANKMHPNIVQYRESFENGSLYIVMDYCEGGDLFKRI---NMQKGLNEDDILDMFOICLAKHV 112
DB 83 IIMKLNHPNVLRLYDWEENTNDYLVLEVAKEGELFNLL-VERGL-PEHAIAFFNQL 140
QY 113 CLAKHVHDRKILHRDIKSONIFLTKDGTVOGLDFGIARVNSTVE--IARCTIGTPYLSPECEN 177
DB 141 IIGVSYCHALGIVARDKLPENLLDHKNYKIKADPGMA-ALTEBKLLETSGSHYAP 199
QY 173 EICENKPYNN-KSDIWAQCVLELCTLKHAFA--EAGSKMKNLVKTIISGSPVSLHYTDLRSLYSQ 237
DB 200 EIVSGIYGFASDVWSGCVILFALTGLRPLPDEEDGNIRTLILVQGEFEMSDDEIS 259
QY 229 YDKLSVQLKFRNRPDRPSVNSILEKGFIAKRIEKL 266
DB 260 READDLIRKLILYDLPFRKRIKTRDLKHPLOKYPISIRDSKSRIGLPDRDYTLPLSSNS 319
QY 275 FCLKTSK-----FSGSPTIPAKRPASGNSISVPAQIKTRPAK-----YGIPLAYKK 323
DB 320 SIDAATLQNLVLLMGRDEGIR-----EKLREPGANAEKTLVYALYRFXC 365

```

QY 324 YGDKRLHEKKPLOKHOANO---TEPKVNTGEEPRK-----ISEBARKRLFEIEK 375
 DB 366 DTOKELIKQOQVKKROSISVSVPKSVITTPQRRRRRELLISVTSRKKPISEF-NKFT 424
 QY 376 KOKDOITSLKAKOAKROKEREKLERINRAREGOMRNVLASAGSGVAKAPFGSGGTAPS 435
 DB 425 ASSASSSMLTPGSSKR-----LSKNPKSKKRLSTIVNOSPTPA 464
 QY 436 SPSSRGQYEHYHAFQMOOORAEADNEAKKREITGRGP-----EROKQLAVE 485
 DB 465 SRNKKAS-----VINWKNKRKASIPSTTKKKRSRSRKRSLIPSKRRESVTTKLMST 519
 QY 486 RAKOVE-----EFLORK-REAMONKARAEQHMVYLAHQIRLONNEQOQKAKLGEKK 540
 DB 520 YAKLAEDDDMEYIEKETKRTSSNFATLIDEIFEYKEQIR-----KEKEELERKVRK 574
 QY 541 EANHSEGESEADMRKRTESLKAHANAAVLEQOLERRKREKAYEREKKWEEHLVA 600
 DB 575 -----AREELERRRROKEEKE-----RAKLEEKEDKROEEL-----KKQIE----- 613
 QY 601 KGVKSSDVSPPIGOHETGSGSPSKOAMRSVISTALKVEGVDSLTDTRETISEMOKTNN 660
 DB 614 --IDISLDEOELSKH-----KEELDGNIRKISAPMENEKIN 650
 QY 661 AISSKRE-TLRLNENLKAOEDKQMONLSDTEINVEDAKEREKESVSDRKKWENG 719
 DB 651 HLEVDIDNLRFRNFSLOT----- 670
 QY 720 GGLVPLDELTLDTSTFTRHVTGVEVILGPNCSPRANGKSPDVSULKILGEALQLQ 779
 DB 671 -----PVSRIDGIMFS-----SPTBEV----- 688
 QY 780 TELLENTTTRSEISPEGEKRYPLITGEKKVOCI--SHEINSAIVDSVPVTKPSESEAS 837
 DB 689 -----SPEPKRENERLTTEKILFTIRSKFGLSSFNIDELKSMKPEPSII 738
 QY 838 POKSLKLEGNLEPPDLETEILOEBSGTNKBESLCTITDWMSE----- 882
 DB 739 APOFLSEERVSDSNDGVESLILPDGKNGVGLKDSITTPAPVGDGLKRISEIRVPOFT 798
 QY 883 EKETRETOSADRTTIOENEVEDGVASTYVDLSIHIHPGNDOSHCKCVDQSVQDEPP 942
 DB 799 RKSRRFSESNRKRLSVLSMTSTKESFTNVDILKNGND--VNNQSGRIPTPRGADSEF 856
 QY 943 FHKVYHSEHLNVPVOVSQCSPESEFARSHSLPKNKKKNSLILGLSTGLPDANPK 1002
 DB 857 LFEVNEE-----AETGNSNDRILYDVGDTIIKKSALKLNFADRENGSNEAK 906
 QY 1003 MLRTCSLPDLSKLEFRLMDVPTVGVKRONLEIDIKENIKESPS--DSEDIVFETD 1059
 DB 907 OTDNLHLPILP-----PLNGD-----NELKQNSQEGDOAHPIKSMIPESGS 949
 QY 1060 TDLOELQASMEOLLREOPGEYSSEESVILKN---SDVPTANGDVADDEDONPSSEAL 1116
 DB 950 SHTKREKENEKEKRP-BOHKOEBOEKREKVVYDMDPPLNKSQVKIRKRNAGSOA-- 1006
 QY 1117 NEWMSONGDEIASECECDSVFNHLELHLEJEMGEKFEFVEYERKKAHDEDEDNI 1176
 DB 1007 -----KDKSK-----DHLEKHODKNTAIGNGSEFRKFSK-----SDKTM 1042
 QY 1177 EICSKIVONILGN 1189
 DB 1043 ELKATLSAKOLF 1055
 RESULT 11
 NIM1_NIMUCR STANDARD: PRT; 779 AA.
 AC P48479;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE G2-specific protein kinase nim-1 (EC 2.7.1.-).

GN NIM-1.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariatales; Sordariaceae; Neurospora.
 OC NCBI-Taxid-5141;
 RX SEQUENCE FROM N.A.
 RA MEDLINE-95355415; PubMed-7629122;
 RA Pu R.T.'s A.;
 RA Pu R.T.'s X.G.; Wu L.; Yierula J.; O'Donnell K.; Ye X.S.;
 RT Annotation of a functional homolog of the cell cycle-specific NIMA
 RT protein kinase of *Aspergillus nidulans* and functional analysis of
 RT conserved residues.
 RT J Biol Chem. 270:18110-18116(1995).
 CC -!- FUNCTION: PROTEIN KINASE THAT PLAYS AN IMPORTANT ROLE IN MITOTIC
 CC REGULATION.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- DEVELOPMENTAL STAGE: ACCUMULATES WHEN CELLS ARE ARRESTED IN G2;
 CC DEGRADED AS CELLS TRAVERSE MITOSIS (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIMA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by isb-sib.ch/announce/
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L42573; AAA80145.1; -.
 DR HSSP: P11362; IFCG.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002290; Ser_Thr_Pkinase.
 DR Pfam: PF00069; Kinase_2.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE-NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_DOM_1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM_1.
 KW transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
 KW Nuclear protein; Phosphorylation; Cell cycle; Cell division.
 FT DOMAIN 7 250
 FT NP_BIND 13 21
 FT BINDING 36 26
 FT ACT_SITE 161 181
 FT DOM_RES 194 759
 FT DOWIN 751
 FT MOVIN 751
 FT SEQUENCE 779 AA; 86079 MW; 7C174925A25EB9B9 CRC64;
 Query Match 8.0%; Score 499.5; DB 1; Length 779;
 Best local Similarity 22.0%; Pred. No. 3.8e-12;
 Matches 211; Conservative 129; Mismatches 326; Indels 291; Gaps 31;
 QY 2 EKYVRLQKIBSGSGKAIIVKSPEDGROYVIEINISRWSSKREEREREVAVLANKKRP 61
 DB 5 DTELLEKIRIGSGGIRKVRKRAKDMILCTREISYLYKMSQKEREOLJAEPSILSTLRHP 64
 QY 62 NITVOY--RESEENSGSIYIYDCEGDLFKRL-NAQGVLFQEOQILDM--FVQICLAL 116
 DB 65 NIVGYHHEHAKQODHLHYMEYCGNGDLGRVIRNLIKINOVAEESFV-WEIFQALVTAL 123
 QY 117 KHVH-----DKRLHROIKSONIPLTKDGVOLGDFGIA 150
 DB 124 YRCHGVDPPEVGVTVLGLSTAPKPPSGCMTILHRDKLPENVLGDNVSKLGDGELS 183
 QY 151 RYLNSTVELARTGIGTPYVLSPEICEENKPYNNKSDIWAIGVLYELCTLKAPEFASGMN 210
 DB 184 KVMQSH-DFASTYVGTPEYMSPEICAEKTYLSDSDVSLGCTIYELCARPEPFNAKTHYQ 242
 QY 211 LVLTITSGSPVSLHYEYDLSIVLSVOLFKNRDRSPVNSILKEGFLAKRIEFLSQL 270
 DB 243 LVQKIEKGIAPLBSVYSGELFATIKDCLRVNDRDRPTATLL----- 285

```

OY 271 IAEFFCTKTESKSGSOPIPAKRPASGONSISVMPAQKITKPAKXGCIPLAKKYGDKLH 330
DB 286 -----NLTIV-----RLKR 294
OY 331 EKKPLOKHOAHOPTPEKRVNTEGERKISEEAKRRKRLFEIEKEKKQKODIISIMKAEOM 390
DB 295 KEKEVEEFTHTLTKKEETLN-----KRIE-----LDSKLSALETER-----SSIRAE-- 337
OY 391 KROEKEKLEKINRAREOGARNVLISAGSGEVAKAPLGSGETIAPSFSRSGOYEHYIAIF 450
DB 338 -----IDASLRRE-----EVKA-----RLEI 354
OY 451 DOMOQOAEADNEAKMKREI-----YGRG-----LPEROGOLAVERAKOYEETFORK 497
DB 355 DRIVAOEITSLQKFEEOEVOAVERAELORHGRGPFNSHGCGSFSSTAATLVSDY----- 410
OY 498 REMOKKAAEGBHAYLARLOIRLQNEERQOITKANL-----RGEKKKANHSEGOEGS 551
DB 411 -NLSSVSGGDDPFSTDDIDIDISTAESTDGSDITKRIPTPHRAQOTYSAPAESEVLGT 468
OY 552 -----EADMRKKIESLKAHANARAAVLEKQLEKRR-----KEAYEREKKEVEHLAKGV 603
DB 469 PMDIEAASPPTIASLSL-SPRKAALAKAPTNPRIEFGSEPTSTDKSNWEVPRETEMI 527
OY 604 KSDVS-----PILQHEHG-----GSPSKQOMRSV 629
DB 528 DSGDESEAEALVPSPKRTKSKNPSTVTYTRRSPSLNSOONSVLPIHGLSKOTLAR 587
OY 630 ISVTALKEGVND-----SLTDTRETSEMOKTNNAISKREILRLNENLKAEDEKG 684
DB 588 SKTVAGSISGIDHPLRSAPSPSLRDKKPSPTRLSLRISPSVGVGRRLANNINNSNG 647
OY 685 MONLSDTEINVDHDAKEHEKESVSDRKKMAGQOLVPLDELTLDTSTFSTERHTVG 744
DB 648 SDASSVTYSNITVTRGLKRMSTCD-----ESSFSQOQ----- 682
OY 745 EVIKLPGSPRRAMKSPPTSDVLIKIGEAELDLOLELLENTINSEISPE-----GE 797
DB 683 -----NNOPOOSLPAPP-----LKKIG-----LMAAKNIGSSLVEIHOARAGR 723
OY 798 KYKPLINGEKVOCI-SHEINSAIYDVPETKSPFSSPOMLAKIEGNLEPDD 853
DB 724 PISAIISNEAKLAFKREHATIASAVDS-----SSSSSSSSQSQSLPTPRSPQPTD 775

RESULT 12
STIO_MOUSE
ID STIO_MOUSE STANDARD: PRT: 966 AA.
DC 025098;
DB 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Serine/threonine-protein kinase 10 (EC 2.7.1.37) (Lymphocyte-oriented
GN SKI0.0 LOK.
OS Mus musculus (Mouse).
OC Fukuyama, Y.; Choudhary, V.; Rodentia; Sciurognathi; Muridae; Mus.
OC Mammalia; Eutelestomi;
OC NCBI-TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=thymus;
RX MEDLINE=97426413; PubMed=9278426;
RA Kurumochi S, Moriguchi T., Kuida K., Endo J., Semba K., Nishida E.,
RT FLOK is a novel mouse STE20-like protein kinase that is expressed
RT predominantly in lymphocytes;
RT J. Biol. Chem. 272:32678-32684(1997).
CC -1- FUNCTION: CAN ACT ON SUBSTRATES SUCH AS MYELIN BASIC PROTEIN AND
CC HISTONE IIA ON SERINE AND THREONINE RESIDUES.
CC -1- CATALYTIC ACTIVITY: ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN LYMPHOID ORGANS
CC SUCH AS SPLEEN, THYMUS, AND BONE MARROW.

```

```

CC -1- PTM: AUTOPHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC STE20 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL databank.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is not used for
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: D89728; BAA24073.1;
CC HSBP; P24941; ICRP.
CC MGD; MGI:1099439; SKI0.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; Pkinase; 1.
CC SMART: SM00220; S_TKc; 1.
CC PROSITE: PS00107; PROTEIN KINASE, ATP; 1.
CC PROSITE: PS00108; PROTEIN KINASE, ST; 1.
CC PROSITE: PS00111; PROTEIN KINASE, DOM; 1.
CC Kinase, serine/threonine-protein kinase; ATP-binding;
CC phosphorylation; Coiled coil.
CC DOMAIN 588 294
CC FT DOMAIN 588 294
CC FT BIND 749 883
CC FT BINDING 42 25
CC ACT SITE 157 157
CC ACT SITE 157 157
CC SEQUENCE 966 AA; 111992 MW; 7115EAC01032BF94 CRC64;

Query Match 7.58; Score 468.5; DB 1; Length 966;
Best local similarity 21.88; Pred. No. 6.6e-11;
Matches 179; Conservative 146; Mismatches 294; Indels 201; Gaps 28;

OY 9 KIGESFGKALVKTEDGRQYIVKEINISMSSKEREESREAVLANKKHPNIVYRE 68
DB 41 ELGDGAFGKVKAKKKEKGAALAAKVIETK--SEELIEDYIETELACDHPYVKLG 98
OY 69 SPEENGSLYIWDYCEGGDLFK-RINQKGVILQEOQILDMVOICLAKHVHDKRIIHR 127
DB 99 AYYDGKLMIMEPCPGAVDAIMLELDRGL--TEQIYQVCKOMLEALNLFGRRIIHR 156
OY 128 DIKSONIFLTGKGTVOGDFGIARVNSTYELARCTIGPYIISPT--CE--NKPNN 182
DB 157 DLKGNVLMTELEGDRIADFGVSANKLTLQKRDSPICGTPYMAEVLCTMKDAPYDY 216
OY 183 KSDIWAIGCVYEELCTKHAPEAGSMKNLVKITISGPPPVSL---HSYDLRSIVSOLF 239
DB 217 KADIVSLGITLIEAQAQIEPHNELNPNRVLLK-IAKSDPPTLLTPSKSVFEPDRLKAL 275
OY 240 KRNPRDRSVNSILKEGFIARKIEKFLSPQLIAE----- 273
DB 276 DKNPETRPSAOLLOHGFPSRVSNKRLRELVAEKAEMEIEEDGREDEEDAVDAVP 335
OY 274 -----EFCLEKTESKSGSOPIPAKRP-----ASGONSI-SVMPAQK 307
DB 336 PLVNHOTDSANVTOPSLDSNKLLODSSTPLPSPQOEPVNGPCSPGDGPIQTTSPADG 395
OY 308 ITRPAKYGIPLAKYKYGDKLHEKKPLOKHOAHOPTPEKRVNTEGERKISEEAKRR 367
DB 396 LSKNDNDLAKVPLRK-----SRPLSMARLIMDEKEKIPPODER--PSPAASKSQ 444
OY 368 LETIEKREKOKDOIISLMKAEMKROEKERLERINRAREOGARNVLISAGSGEVAKAPLG 427
DB 445 -----RANO-SRPNSSALETI-----GGEA-----LT 465
OY 428 SGGTIAPSFSRSGOYEHYIAIFDOMOQOAEAD--NEAKMKREIYGNGLP-----EROKG 480
DB 466 NGGLELPSVY-----PSHSKRASDCSNLSTESMDYGTSLADSLNKRERG 512

```


QY 686 QNLSDFEINVEDAKEHEK 707
 Db 760 QOLKDOYFLORHLLRKHHEK 781

RESULT 14
 KKRL-YEAST STANDARD; PRI: 1518 AA.

AC P34244;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Probable serine/threonine-protein kinase YKL101w (EC 2.7.1.-).
 GN YKL101w OR YKL453.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OX Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 RN YKL101w OR YKL453;
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE-94078677; PubMed-8256524;
 RA Pallier C., Valens M., Puzos V., Fukuhara H., Cheret G., Sor F.,
 RA Belcien-Fukuhara M.,
 RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI
 RT physically localizes the MBI gene and reveals eight new open reading
 RT frames, including a homologue of the KINI/KIN2 and SNF1 protein
 RT Yeast 9:1149-1155(1993).
 RL -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1 NIM1 SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).

DR EMBL: X71133; CAA50456.1;
 DR EMBL: Z28101; CAA81941.1;
 DR PIR: S37928; S37928.
 DR PIR: S39084; S39084.
 DR HSP: P13362; 1FGK.
 DR SGD: S0001584; YKL101w
 DR Interpro: IPR000719; Euk-kinase.
 DR Interpro: IPR002290; Ser-thr-kinase.
 DR Pfam: PF00069; kinase_1.
 DR SMART: SM00220; S_TKc_1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP. 1.
 DR PROSITE: PS00108; PROTEIN KINASE ST. 1.
 DR PROSITE: PS00111; PROTEIN KINASE DOM. 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding.
 FT DOMAIN 81 369 PROTEIN KINASE
 FT NF-BIND 87 95 ATP (BY SIMILARITY).
 FT BINDING 110 110 ATP (BY SIMILARITY).
 FT ACT_SITE 239 239 BY SIMILARITY.
 FT SITE 239 239 BY SIMILARITY.
 SEQUENCE 1518 AA; 169592 MW; 803f84f7531241DD CRC64;

Query Match 7.2%; Score 452; DR 1; Length 1518;
 Best local similarity 20.2%; Pred. No 4,3e-10;
 Matches 260; Conservative 199; Mismatches 449; Indels 380; Gaps 50;

QY 10 IGGSGGKALVKESTEDGROYIKETI-----NISRMSK 43
 Db 87 LGSGSGGKALVKESTEDGROYIKETI-----NISRMSK 43

QY 44 ERRES-----RREAVAVLANKKHPNIYQYRESFEENSATYIVNDYCEGDJLRIN 93
 Db 147 SRHSNHSQTNPNYGIERTIVIMKILSHTNVAFLEWENNSSELYLVLEIVDGGELDYI- 205

QY 94 AOKGVLFQEDQIDLDFVQICLALKHVDRKILHRDIKSONIF--TKDGVQGLDPCRIARV 152
 Db 206 VSKGRL-PEREAHIFKQIVGVSYCHSFNCHRDCLKPENLLDKNNRIKTIADYMA-A 263

QY 153 INSTVELARTCIGTPYULSPICENKRPYN-NKSDYALGCVLELCTLHAFAEASMKL 211
 Db 264 IELPNKLTSCGSPHYASPEIVMGREYHGGSDVMSCGIVFALLGLHPNDONIKKL 323

QY 212 VKIITSGSFPPVSLHYSDLSLVSOLFKNPNDRPSVNSIIEKGTAK----- 260
 Db 324 LKVOGSKGYQWPS-NLSSEARDLISKILVIDPEKRITTOEILKHLPIKKYDLPVNVKLR 382

QY 261 -RIEKLPSLIDLEFCLTFKFSQOPIPARFPASGNSISVPAQKITPAKAGI- 317
 Db 383 KMRDNNARCSNDIHLNNVS-----PSIVTLHSGEIDISILRSIQI-LMGVS 433

QY 318 -----PLAYKKYGDKKLHEKKPLQKHKAQHQPPEKRVNTGGERKISEAARKRLFEI 371
 Db 434 RELITAKLQPMSEKLFYSLLO-YKORHSI-SLSSSENKKSATSSVNERIERYA 490

QY 372 EKE-----KKQDQIISLAKADQMKROEKERLERINRAREGQWNVLSAGSGEVKAPF 425
 Db 491 SKTANNTGLSENNDVKTLSLE-----HSEDTSTVN-----ONNAITGVNTEINAPV 539

QY 426 LGSGGTTAPSSFSRQYEHYNAIFDQMOQRAEDNEAKWKREIYGRQLPERQGLAVE 485
 Db 540 LA-----OKSOPSIN 549

QY 486 RAKQVEEFLQKKREAMQNAKAGHMVYL-----ARLQIRLQNFENRQQIKAK 534
 Db 550 TLDQPE-----SDKAEDA-VTLPAIPFNASSIRFNSTYSSISRSRSLR 597

QY 535 LREKKFANHSQEQEESQADN-----RRKIESLKAHANAARVLEQDERRKE 585
 Db 598 LMSRSLKSLSTRKREIVHNDENPRLDQPKSPRSLSRAIHAASPKRSJHKSIRKNIA 657

QY 586 AYREKKVVEHLYAKGVSSVSPPLQOHEGGSPPSKQOMRSYISVTA--LKEVGV- 641
 Db 658 ATVAARITQ-----NSAKRSLYSQSTSKSLNLDILYF 694

QY 642 DSSLTDTRETSREMOKT-----NNAISSKRETLRLNENLKAQED 681
 Db 695 DDLPLSKPASNKSEPHSLSDPEILCDQILFGNALD--RLLEEEDNEKERDT 751

QY 682 EKGQMN-----LSDTFEINVEDAKEH-----EKEKSVSDRKKWEAGQVLIP- 726
 Db 752 OROFQNDTKSSADDTTIGSVSTKNKEBPETYLEKQNMSTKPSSENGSLSPFIRE 811

QY 727 DELTLDTSFTSTERTVGEVYIKLPNGSPRAAKSKTIDSVYLKILGAELOJOTELLENT 786
 Db 812 KENTLSSSYLERQK-----PKNALSDITNSNKNKQOEGRIEKTQ-- 854

QY 787 TIRSEISPEGEKRYKPLINGEKKYOCISHEINPSAIVSPETYSPPSEASPMQSLKLG 846
 Db 855 --RQLOLQKNDRPSPL-----KPIQHQ-----ELRVSLPNDQKPSLSDPERR 896

QY 847 NLEPPDDELETEL-----QEPSC--TNKDESI-----PCTIDVVISER--K 884
 Db 897 NISQPVNSKVESLLOGLKFKKEPASHWTHRGSLFMSHVEDEKPKASDVSISSSIYPL 956

QY 885 EKEFQASRDTTIOENESD-----GVSSVYDLSDDHIERPNDOSHKCDVD 934
 Db 957 TTVATISSRDPVLAESSITQKPMLSLSPSSPLNTSMFKNLSDIADDDG----- 1006

QY 935 KSVQPEPFHVVHSEHLNVLPQVQSVQCSPEESFAFRSHS--HLPKNNKKNLLGL 991
 Db 1007 -----KHLS-VPOQNS--RNVAMSHPLRQKQAKISLITRSNINAMNLVKR 1048

QY 992 STGLFDANNPMLRTCSLPLSLKFLTMDVPLVPGVQRQNDLEIDEINDENKREPSQSE 1051
 Db 1049 NOG-----SPGSYLSNDLDGIDSM--TFAMEIPI-----NTTQAO 1081


```

QY 1052 DIVFEETDIDQEQAS-----MOLLREQGESESESVKNSKSPFANGCQVDA 1104
Db 1082 AIOAMNDNDNNKNTSKKASFTKEKVIKS--AAVTSKEKPEPNDSPNTIPPTTPNTY 1139
QY 1105 DE-----DDNESSESALNEBWHSDNSD 1126
Db 1140 DEKAINIFEDAFSDGSLNTS--SSESD 1165

RESULT 15
KCC4_YEAST STANDARD; PRT; 915 AA.
AC P25389; P87005;
DT 01-MAY-1992 (Rel. 22, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Probable serine/threonine-protein kinase YJC24W (bc 2.7.1.-).
YC1024W OR YJC24W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota, Fungi, Ascomycota, Saccharomycotina, Saccharomycetes;
OC Saccharomycetales, Saccharomycetaceae, Saccharomyces.
OC NCBI_taxid=4932;
NM
RN
RA SEQUENCE OF 567-915 FROM N.A.
RA Oliver S.G., Awmar R., Brown A., Gent M.E., Indge K.J., James C.M.,
RA Staveja L.I.;
RL submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
[2]
RN
RP SEQUENCE OF 567-915 FROM N.A.
RP Gromacka R.;
RL submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
[3]
RN
RP REVISIONS.
RP Gromacka R.;
RL submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NML SUBFAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch)
CC -----
CC EMBL; X59720; CAA42361.1; -
CC PIR; S19351; S19351.
CC HSSP; P00518; IPRK.
CC SGD; S0000529; YC1024W.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase.1.
CC PRINTS; PR00109; TYRKINASE.
CC SMART; SM00220; S_TKC.1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP.1.
CC PROSITE; PS00108; PROTEIN_KINASE_SF.1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM.1.
CC Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding.
FT DOMAIN 21 285 PROTEIN KINASE.
FT NP BIND 27 35 ATP (BY SIMILARITY).
FT BINDING 50 50 ATP (BY SIMILARITY).
FT ACT SITE 152 152 BY SIMILARITY.
FT FT ACT SITE 152 152 BY SIMILARITY.
SQ SEQUENCE 915 AA; 102688 MW; BFE0108CA43AC181 CRC64;

```

Wed May 15 14:14:05 2002

us-09-783-320-4.rsp

Page 14

Db 874 -----ENTEDGAEKGTESKTAASLSPDDUK 900

Search completed: May 15, 2002, 08:02:27
Job time: 268 sec

OM of: US-09-783-320-4 to: EST.* out_format: pfs

Date: May 15, 2002 10:47 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp
-O=cgnt2.1/USP/USP0183320/runat.14052002.105145.1592/app-query.fasta.1.1290
-DB=EST -OEMT=fastap -SUFFIX=est -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPE=0.000 -LOOPEXT=0.500 -GAPOP=6.000
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blom62 -TRANS=blom60 cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MILEN=0 -MAXLEN=200000000 -USER=US09783320 -CGNT2.1.8913
-NCPU=6 -ICPU=3 -LONGLOG -DEV=TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPHY -MAIT -THREADS=1

Search information block:

Query: US-09-783-320-4
Database: EST.*
Database sequences: 13736207
Database length: -1841457050
Search time (sec): 3581.300000

Score	List:	Strid	Orig	ZScore	Escore	Len	Document
90.0	BC015147	2038.00	2716.19	3.7e-143	1628	1	BC015147 Homo sapiens, clone
90.0	BC015147	+2020.00	2712.62	7.6e-142	1515	1	AK009484 Mus musculus adult
90.0	BC015147	+1906.50	2559.24	2.6e-133	1515	1	AK009484 Mus musculus adult
90.0	BC015147	-1177.00	1581.67	7.5e-79	692	1	AK009484 Mus musculus adult
90.0	BC015147	-1171.00	1571.83	2.6e-78	816	1	AK009484 Mus musculus adult
90.0	BC015147	+1092.50	1467.08	1.8e-72	719	1	BC015147 Homo sapiens, clone
90.0	BC015147	+1075.00	1441.72	4.7e-71	847	1	BC015147 Homo sapiens, clone
90.0	BC015147	+1057.00	1430.32	7.2e-70	641	1	AV721713 AV721713 Hb
90.0	BC015147	+1023.00	1374.35	2.6e-67	642	1	AV721713 AV721713 Hb
90.0	BC015147	-1012.00	1364.70	9.1e-67	654	1	AV721713 AV721713 Hb
90.0	BC015147	-1006.50	1349.18	6.6e-66	548	1	AV721713 AV721713 Hb
90.0	BC015147	-931.00	1093.69	1.1e-51	477	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	-813.00	1087.99	3.1e-50	544	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+795.00	1087.99	3.1e-50	544	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+767.00	1032.12	3.0e-48	451	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	-766.50	1018.68	5.3e-47	463	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	-746.00	1019.37	1.0e-47	702	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+746.00	1004.15	1.1e-46	488	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	-741.00	1002.92	5.6e-46	761	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	-719.00	967.71	3.0e-44	432	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	-719.00	960.44	3.0e-44	454	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	-709.50	950.44	1.1e-43	659	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+705.00	943.35	2.7e-43	774	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+699.00	939.86	4.2e-42	654	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+695.00	930.93	1.3e-42	654	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+691.00	920.93	2.4e-42	793	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+666.00	889.72	2.6e-40	630	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+662.50	887.40	3.5e-40	630	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+654.00	888.97	2.9e-40	604	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+616.00	822.76	1.4e-36	748	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+614.00	822.76	1.4e-36	643	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+609.50	817.24	9.8e-36	548	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+603.50	807.52	2.8e-35	639	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+602.00	805.51	1.3e-35	638	1	AA907081 AA907081 NIH-MGC
90.0	BC015147	+599.50	801.62	2.1e-35	670	1	AA907081 AA907081 NIH-MGC

seq_name: gb_hic:BC015147
seq_documentation_block:
LOCUS BC015147 1628 bp mRNA linear HTG 04-OCT-2001
DEFINITION Homo sapiens, clone IMAGE:4063881, mRNA.
ACCESSION BC015147
VERSION BC015147.1 GI:15929426
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 1628)

REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: GENE TECH Laboratories, Inc.
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: nadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahy, Stephanie Ford, Julia
Greene, Mark Kettman and Anuradha Madan

FEATURES
Source
This clone has the following problem: frame shifted.
1. 1628
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4063881"
/tissue="Bone marrow, chronic myelogenous leukemia"
/clone_lib="NIH-MGC_54"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

BASE COUNT
549 a 287 c 360 g 432 t
ORIGIN

alignment_scores:
Quality: 2038.00 Length: 400
Ratio: 5.108 Gaps: 0
Percent Similarity: 99.750 Percent Identity: 98.750

Align seg 1/1 to: BC015147 from: 1 to: 1628
1 MetGluTyrValAlaGluGlnIleGlyGluGlySerPheGly 17
|||||
419 ATGGAGGATGTTAGCTACGACATGGCAAGCTTCATTGGA 468
|||||
17 salalaleuvallysserthrguaspglaragintyvalilley 34
|||||

469 AGCCATCTCTGTTAAATCTACAGAGATGGACAGACAGATGTTATCAAGG 518
 34 IUILEASIIIESEFARIGMETSERIYSLUARGIUGIUGIUSERARFAR 50
 1419 AAATTAACATCTCAAGAAATGCCAGTAAGAAAGAAACAAATCAAGAGCA 568
 51 GIUVALIAVALLIEUALIASENMELIYSHISPROASIIIEVALGINTYAR 67
 559 GAAGTGGAGTATTGGCAACACATGACATCAATATGTCAGATAG 618
 67 gLUSeRPhEGlUGIUSnGLYSeRleuTYrIIeValMeTAsPYrCYSG 84
 619 AGAATCATTTGAGAAAAATGGCTCTCTACACAGTAAGATGATCTG 668
 84 IUGIUGIYASPLUEUPHELYSARGIIIEASNAIAGIINISGLIYALUEUPHE 100
 669 AGCGAGGGGATCTGTTAAGCGAATTAATGCTCAGAAAAGCGCTTTGTTT 718
 101 GINUGIUSPGLIIELEUASPTRPHEVALGIIIECYSELEUALIEULY 117
 719 CAAGGAGTACGATTTTGGACCTGTTTGAACAGATATGTTGGCCCTGA 768
 117 SHIVAIHISAPARGIYSLIELEUHSARGSPILIELYSERGIIAN 134
 769 ACATGTACATGATAGAAAAATTTCTCATGACACATTAATTCACAGACH 818
 134 IEPHELEUTHIRLYSASPGLYTHIRVALIGIUGIUGIYASPHEGIIIEALA 150
 819 TATTTTACTAAAGATGAGAACAGTACACTGAGATTTTGGAAATGCT 868
 151 ARGVALIEUASNSERTHRVALIGIUEUALIARGTHIRCYSLIEGLYTHIR 167
 869 AGAGTCTTAATAGTACTGTAGAGCTGGCCGAACTTGATGGAGACCC 918
 167 OTYTYRILEUSERPROGIUIIECYSGIUSnGLYSPROTARYANASnLYSS 184
 919 ATACTACTGTGCACCTGAAATCTGTGAAAAAACCCTTACAAATATAAAA 968
 184 ERASPILETFRPALALEUGIYCYSEVALLEUTYRIGIUEUCYSTRHEULYS 200
 969 GTGACATTTGGGCTCTGGGCTGTCTCTTATAGCTGTACTACTTAA 1018
 201 HISIAPHEGIUAIAGIYSERMELIYASnLEUVALIEULYSIIIESE 217
 1019 CATGCTTTTGAAGCTGGCAGTATGAAAAACCTGGTACTGACATATATATC 1068
 217 RGIYSERPHEPROVALSERIEUHSIYSETYRAsPLEUARGSERL 234
 1069 TGGATCTTTTCACCTGTGTCTTTCATTTCTATGTGATGCCAGATT 1118
 234 EUVALISERGIULEUPHELYSARGASnPROARGSPARPROSERVALAsn 250
 1119 TGGGTGTCACTATTATTAAGAAATCCAGAGGTAGACCACTGCTCAAC 1168
 251 SEITIELEUGIUGIYPHEIIIEALIASARGIIIEGLIULYSPhEUSE 267
 1169 TCATATATTGGAGAAAGCTTTTATAGCCAAACCATGAAATCTCTCTC 1218
 267 RPROGIULEUIIEALIGIUGIUPHECYSEULYSSTRPHASERLYSPHEG 284
 1219 TCTCTAGCTATTTCACAGAAATTTTGTCTAAAAACATTTTCGAAGTTTG 1268
 284 IYSERGIIINPROIIEPROIALISARGPROIASERGIYIINAsnSERIIE 300
 1269 GATCAGACCCATATACAGCTAAAGACCACTTCAGAGACAAAACGCGATT 1318
 301 SERVALIETPROIAGIUGIULYSIIIEHIRLYSPROIALIASLYTYGIIY 317
 1319 TCTGTATTAGCTGCTCAGAAAAATTACAAAGCCGCGCAATATATGGAAT 1368
 317 EPROLEUALATYRILYSLYSTYRGIYASPIYSLYSLEUHSIGIULYS 334
 1369 ACCCTTACCATATTAAGAAATATGAGATTAATAAAATTTACAGAAAGAAAC 1418

334 roleUGIULYSIIIEHISGLIUALIIEHISGLIUPHROGIULYSARGVALAsn 350
 1419 CACTCCAAAAACATTAACAGCCCATCAATCAACAGAGAGAGGAGAT 1468
 351 THRGIGIUGIUGIARGARGIYSLIESERGIIUALIHALIARGLYSARG 367
 1469 ACTGGAGAAAGAAAGAGAAATATCTGAGGAAGCAGCAAGAAAGAG 1518
 367 gLEUGIUPHEIIIEGLIULYSGLIULYSGLIULYSASPGLIIEISERL 384
 1519 GCTGAATTTATTTAAAAAGAAAAAGAAAGGATGAGATTTATTTACTT 1568
 384 eUMELIYASIIAGIUGIUMELIYSARGIUGIUGIUGIUGIUGIUGI 400
 1569 TAATGAGGCTGAAACAAATGAAAAAGCCAGCAAAAAA 1618
 seq_name: gb_hic:AK010196

seq_documentation_block: 1515 bp mRNA linear HTC 19-JAN-2002
 LOCUS AK010196
 DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310076F08:NIM (never in mitosis gene a)-related expressed kinase 1, full insert sequence.
 ACCESSION AK010196
 VERSION AK010196.1 GI:12845466
 KEYWORDS HTC; GAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone:2310076F08.
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 (sites)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 MEDLINE 10349636
 PUBMED 10349636
 REFERENCE
 2 (sites)
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

TITLE
 JOURNAL 20499374
 MEDLINE 11042139
 PUBMED 11042139
 REFERENCE
 3 (sites)
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsu, T., Taahiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaiguchi, S., Ikegami, T., Kishiyagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

TITLE
 JOURNAL 20530913
 MEDLINE 11076861
 PUBMED 11076861
 REFERENCE
 4 (sites)
 The RIKEN genome Exploration Research Group Phase II Team and the FANOM Consortium
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

5 (bases 1 to 1515)
 Aachari, V., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bulic, G., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiroaka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Iizawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,

seq_name: gb_hc:AK009484

seq. documentation block:	1515 bp	mrna	linear	HTC 19-JAN-2002
LOCUS	AK009484			
DEFINITION	Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310022M10.N1MA (never in mitosis gene a)-related expressed kinase 1, full insert sequence.			
ACCESSION	AK009484			
VERSION	AK009484.1	GI:12844311		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library			
ORGANISM	clone:2310022M10.			
REFERENCE	Mus musculus			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	1 (sites)			
JOURNAL	Carninci, P. and Hayashizaki, Y.			
PUBMED	High-efficiency full-length cDNA cloning			
REFERENCE	Meth. Enzymol. 303, 19-44 (1999)			
AUTHORS	99279253			
TITLE	10349636			
JOURNAL	2 (sites)			
PUBMED	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
REFERENCE	Normalization and subtraction of cap-trapper-selected cDNAs to			
AUTHORS	Prepare full-length cDNA libraries for rapid discovery of new genes			
TITLE	Genome Res. 10 (10), 1617-1630 (2000)			
JOURNAL	20493374			
PUBMED	11042159			
REFERENCE	3 (sites)			
AUTHORS	Shibata, K., Itoh, M., Atzawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,			

AUTHORS
Shibata, K., Itoh, M., Alzawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitanaka, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamanoto, K., Matsumoto, H., Sakaueguchi, S., Ikegami, T., Kashiwagi, K.,
Fujitake, S., Inoue, K., Togawa, Y., Iwata, M., Ohara, E., Watanabe, K.

TITLE	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Murakawa, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
JOURNAL	RIKEN Integrated sequence analysis (RISA) system-384-format
MEDLINE	sequencing pipeline with 384 multiplexillary sequencer
PUBMED	Genome Res. 10 (11), 1757-1771 (2000)
REFERENCE	20530913
AUTHORS	11076861
	4 (sites)
TITLE	The Riken Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
JOURNAL	Functional annotation of a full-length mouse cDNA collection
REFERENCE	Nature 409, 685-690 (2001)
AUTHORS	5 (bases 1 to 1515)
	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,

TITLE Direct Submission
JOURNAL Submitted (10-JUN-2000) Yoshihide Hayashiaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome.res@sc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

COMMENT

Please visit our web site (<http://genome.jsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGCAGACAGAATCCAGACAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapted of sequence [5'-GAGCAGACAGATTTCGAGTTAATAATTAATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOUR.

FEATURES	
source	Location/Qualifiers
	1. .1515
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="MGI:1905247"
	/db_xref="taxon:10090"
	/clone="2310022M10"
	/sex="male"
	/tissue_type="tongue"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="adult"
gene	1. .1515
	/gene="Nek1"
	1. .1515
	/gene="Nek1"
	/note="NIMA (never in mitosis gene a)-related expressed kinase 1"
misc_feature	data source:MGI, source key:MGI:97303, evidence:ISS"
BASE COUNT	512 a 281 c 374 g 348 t
ORIGIN	

alignment_scores: Length: 500
 Quality: 1906.50 Gaps: 5
 Ratio: 4.190
 Percent Similarity: 91.000 Percent Identity: 79.600

alignment_block:
 US-09-783-320-4 x AK009484 ..

Align seg 1/1 to: AK009484 from: 1 to: 1515

```

151 ArgValLeuAsnSerThrValGluLeuAlaArgThrCysTleGlyThr 167
    |||||.....|
29 CGAGTCTTAATAGT...ATGTAGAGTGCGCTCGAATTCATAGGCACTCC 75
    |||||.....|
167 CTATTTLeuSerProGluTleGlyAsnLysProTyrAsnLys 184
    |||||.....|
76 ATACTATTGTAC...TGAATCTGTAACAAAGCCTTATTAACATTAAA 122
    |||||.....|
184 eAspIleTrpAlaLeuGlyCysVal...LeuTyrGluLeuCysThrLeu 199
    |||||.....|
123 GTGACATTAGGCGCTTGGCGCTGCTCTTTATGAGATTGCGTACACTT 172
    |||||.....|
200 LysHisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIle 216
    |||||.....|
173 AAACATCATTTGACCTGAGAAACATGAAACCTGTACTGAAAGATTAAT 222
    |||||.....|
216 eSerGlySerPheProProValSerLeuHisTyrSerTyrAsnLys 233
    |||||.....|
223 CTCGGATCTCTTCTCCAGTGTCTCCACATTACTCTATGATCTCCGCA 272
    |||||.....|
273 eLeuValSerGluLeuPheLysArgAsnProArgAspArgProSerVal 249
    |||||.....|
273 GCTTGTCTCTCACTTATTAAAGAAATCCTAGGATGACCAATCAGTTC 322
    |||||.....|
273 GCTTGTCTCTCACTTATTAAAGAAATCCTAGGATGACCAATCAGTTC 322
    |||||.....|
250 AsnSerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPhe 266
    |||||.....|
333 AACCTCATTTGGAGAAAGSTTTTATAGCTAAACGATGAAAGATTTCT 372
    |||||.....|
266 uSerProGluLeuIleAlaGluLysPheCysLeuLysThrPheSerLys 283
    |||||.....|
373 CTCCTCTAGCTATTTGCAAGAAATTTTGTCTAAAGAACCTTCAAGT 422
    |||||.....|
283 heGlySerGluProIleProAlaLysArgProAlaSerGlyLysSer 299
    |||||.....|
423 TTGGACACAGCTCTCCAGTAAAGACCACTATGACAGAGGTTC 472
    |||||.....|
300 IleSerValMetProAlaGluLysIleThrLysProAlaAlaLysTyr 316
    |||||.....|
473 AGTTCTTTTGTCTCTCCAGTAAATCAAAACCTGCTTAATTAACG 522
    |||||.....|
316 ylleProLeuAlaTyrLysLysTyrGlyAspLysLysIleuHisGluLys 333
    |||||.....|
523 AGGCCTTTAACATATAGAGATGTGAGATTA...AAGTTCTTGAGAAAA 571
    |||||.....|
333 yGProLeuGluLysHisLysGluAlaHisGluThrProGluLysArgVal 349
    |||||.....|
572 AACCAACCCCAAAACATTAACAGAGCCCTCAATTCCTCCGTAAGATATG 621
    |||||.....|
350 AsnThrGlyGluGluArgArgLysLysSerGluGluAlaAlaArgLysAr 366
    |||||.....|
622 AATTCGCGAGAAAGAGAGATTAATGCTGAGAACACCAAA...AAAA 670
    |||||.....|
366 gATLeuGluPheIleGluLysGluLysGlyLysAspGluIleLys 383
    |||||.....|
671 GAGGTTCGATTTATTGAGAAAGAAAGAAAGCAAAAGCTCAG...ATTA 717
    |||||.....|
383 eLeuMetLysAlaGluGluMetLysArgGluGluLysGluArgLeuGlu 399
    |||||.....|
718 GGTTCCTGAGAGCTGAGACATGAACCCCAAGAGACACACGCTTGAG 767
    |||||.....|
400 ArgIleAsnArgAlaArgGluGluLysTyrPargAsnValLeuSerAla 416
    |||||.....|
768 AGGTAATATGGCCAGGAGACAGATGAGAAATGTTTAAAGCCTCG 817
  
```

```

416 yGlySerGlyGluValLysAlaProPheLeuGlySerGlyThrIleA 433
    |||||.....|
818 TGAAGCGGTGAGATTAAGGCTCTTTTGGCATTTGAGGGCGTCTCT 867
    |||||.....|
433 laProSerSerPheSerSerArgGlyGluTyrGluHisTyrHisAlaIle 449
    |||||.....|
868 CTCATCACCCTGCTGCTCTCCAGGCGAGATTAAGCTTTCATTCGAT 917
    |||||.....|
450 PheAspGluMetGluGluAlaArgAlaGluAspAsnGluAlaTyrTrp 466
    |||||.....|
918 TTGACCAATATCAGCGGTAAAGACAGAGATATATAGACAGATGAA 967
    |||||.....|
466 sArgGluIleTyrGlyArgGlyLeuProGluArgGluLysGluLeuA 483
    |||||.....|
968 GGGGCAATCTATGTGCTGCTGCTCCAGAAAGCAAAAGGACACTTAC 1017
    |||||.....|
483 lValGluArgAlaLysGluValGluGluPheLeuGluArgLysArgGlu 499
    |||||.....|
1018 CTGAGAGAGAGCCCAACAGTGAAGATTCCTACAGCCTTAACGAGAA 1067
    |||||.....|
500 AlameGluAsnLysAlaArgAlaGluGluLysMetValTyrLeuAlaAr 516
    |||||.....|
1068 CCTATGCAATTAAGCCCGAGCCGAGAGACACGTGTTTATTTGGCAAG 1117
    |||||.....|
516 gLeuArgGluIleArgLeuGluAsnPheAsnGluArgGluGluLysA 533
    |||||.....|
1118 ACTAGGCAATATAGACTACAAATTTTATGAGCGCCACAGATTAAG 1167
    |||||.....|
533 lAlysLeuArgGlyGluLysGlyGluAlaAsnHisSerGluGluGlu 549
    |||||.....|
1168 CCNAATCTTGAGGAGAAATTAAGAAAGCTGATGATCCAAAGGACAGAA 1217
    |||||.....|
550 GlySerGluGluAlaAspMetArgArgLysLysIleGluSerLeuLysAl 566
    |||||.....|
1218 GCACCTGAGAGACTGACATGAGCTCAAAACATGAGTACTTAAGGC 1267
    |||||.....|
566 hHisAlaAsnAlaArgAlaAlaValLeuLysGluGluGluAlaGlyLys 583
    |||||.....|
1268 GCAAAACAATCCAGCTCTCTCTACTTAAGAAAGACCTGGAGCAAAA 1317
    |||||.....|
583 gLysGluAlaTyrGluArgGluLysLysValTyrGluGluHisLeuVal 599
    |||||.....|
1318 GAAAGGAAGCTTAAGAAAGAAAGAAAGAAAGTATGGAGAAACATTTGG 1367
    |||||.....|
600 AlAlaLysGlyValLysSerSerAspValSerProProLeuGlyGluHis 616
    |||||.....|
1368 GCGAGG...GTAAAGAGCTAGATGTTCTCTGCTCTTGGAACTTCTGA 1414
    |||||.....|
616 uThrGlyGlySerProSerLysGluGluMetArgSerValIleSerValT 633
    |||||.....|
1415 AACAGGTGTTCTCCATCAAGCAGCAGTGAAGCCTGATTTCTGTGA 1464
    |||||.....|
633 hSerAlaLeuLysGluValGlyAlaAspSerSerLeuThrAspThrArg 649
    |||||.....|
1465 CTTCACCTTGAAGAAATGGGCTGAGTGAAGTTAATCACTATCCAG 1514
    |||||.....|

```

seq_name: gb_est1.A1936517

seq_documentation_block:

LOCUS A1936517 692 bp mRNA linear EST 17-DEC-1999

DEFINITION W28Bh08.x1 Soares.NFL.T.GBC.S1.Homo.sapiens.cDNA clone IMAGE:2329503.3' similar to SW.NEKL.MOUSE.P51954

ACCESSION U000000000

VERSION 1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 692)

AUTHORS NC1-CGAP <http://www.ncbi.nlm.nih.gov/nclogap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.
Insert Length: 1721 Std Error: 0.00
Seq primer: -400P from Glibco
High quality sequence stop: 444.

FEATURES

source
1. 692
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2329503"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker. Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung, NBHL19W, testis NHT, and B-cell
NCL-GAP_GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo.

BASE COUNT 187 a 138 c 128 g 238 t 1 others

alignment_scores:
Quality: 1177.00 Length: 230
Ratio: 5.140 Gaps: 0
Percent Similarity: 99.565 Percent Identity: 98.696

alignment_block:
US-09-783-320-4 x A1936517/rev ..

Align seg 1/1 to reverse of: A1936517 from: 1 to: 692

```

142 VALGILNLEUGIYASPPHEGLYLEALIAARVALLIENUNSERTHYVALGI 158
|||||
692 GTACACCTTGAGATTGGAAATGCTAGAGTCTTAATAGTACTGTAGA 643
158 ULEUALIARGTTCYSLIEGLIYTHRPROTYTYTLEUSERPROGLUINLEC 175
|||||
642 GCTGCTCGAACTTCATAGGAGCCCATCTACTCTGCACCTGAAATCT 593
175 YSGIUAASLYSPROTYRASNASNLYSSEASPILETSPALALEUGLYCYA 191
|||||
592 GTGAACCAACACTTACATATATAAGTGACATTTGGCTCGGGGTGT 543
192 VALLEUTYRGULIUCYSTHLEULYSHIALAPHEGLUAGLSERME 208
|||||
542 CTCCTTTAGACCTGTACACTTAACATGCTTTTAACCTGACATTA 493
208 LYSASNLEUVALLEULYSILEIIESEGLYSERPHROPROVALSERL 225
|||||
492 GAAAAACCTGCTAGTGAAGATATATCTGATCTTTCCACCTGTGCTTT 443
442 TGCATTAATCTATGATCTCCGAGTTGGTGTCTCAGTTATTTAAAGA 393
225 EUHISTYRSETYRASPLEUARGSERLEUVALSERGINLEUPHELYSARG 241
|||||
242 AANPROAASPARGPROSERVALASERLIELEUGLIUYSGLYPHELI 258
|||||
392 AATCTTAGAGATAGACATGCTCACTCCATATTGAGAAAGTTTAT 343
258 EALIALYSARGLIEGLIULYSRHELEUSERPROGINLEULIAGLUGLAP 275
|||||

```

```

342 AGCCAAACGATTGAAAAGTTCTCTCTCCACCTATTCACAGAAAT 293
275 hecysleuysthrpserlyspheglyserginproileproalaly 291
|||||
292 TTGCTTAAAAACATTTTGGAGTTGGATCCACACCTATACAGCTAAA 243
292 ATGPROALASERGLYINSERLIESERVALMETPROALAGLINSII 308
242 AGACCACTTCAGACCAAACTCGATTTCTGTATGCTGCTCAGAAAT 193
308 eThrlySPROALALALYSTYRGYLEPROLEUALATYRYSLYSYRG 325
192 TACAAAGCCGCGCTAAATATGGAATACCTTACATATAGAAATATG 143
325 LYSAPLYSLYSLIENHISGLIULYSYSPROLEUGLINSIISGLIALA 341
142 GAGATTAATAATTTACAGAAAGAAACCACTCCAAAAATTAACAGGCC 93
342 HISGLINTHROGLIULYSARGVALASNTHRGLIUGLUGATGATGLYSII 358
92 CATCAAACTCCAGAGAAGAGTAACTGAGAGAAGAAAGAGGAAAT 43
358 eSERGLIUALALALARGLYSARGTGLIUGLUPHELI 371
42 ATCTGAGGAAGCAGCAAGAAAGANGACTGGAATTAAT 3
seq_name: gb_est1:A1816275

```

seq_documentation_block:

LOCUS A1816275 816 bp mRNA linear EST 09-JUL-1999
DEFINITION au6d12.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone
IMAGE:2517815 3' similar to SW:NEK1.MOUSE P51954
SERINE/THREONINE-PROTEIN KINASE NEK1; mRNA sequence.

ACCESSION A1816275
VERSION A1816275.1 GI:5431821
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 816)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
Kritzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Schellenberg, R., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wyllie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST project
Unpublished (1997)

TITLE
JOURNAL
COMMENT
Other ESTs: au6d12.y1
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -400P from Glibco
High quality sequence stop: 424.

FEATURES

source
1. 816
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2517815"
/clone_lib="Schneider fetal brain 00004"
/sex="male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site_1: SbfI; Site_2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor

alignment_block:
US-09-783-320-4 x BG706222

Align seg 1/1 to: BG706222 from: 1 to: 719

```

850 GUPPOASAPSPleuGluThrgluLeuGluGluProsergLythras 866
7 GAACCTGATTTGGAAAGAAATCTCMAGAGCCAGTCGAAACAA 56
866 nlyaspLuserLeuProcystrilietrraspaValTrrileserGlu 883
57 CAAAGATAGAGCTTCCGCTGCTATTCGATGATGATGATGAG 106
883 lulyGluThrgluThrgluInserAlaAspargIlethrlleGlu 899
107 AAAAAAGAAAGAAAGAACTAGTCGAGATAGGATCACCATTGAG 156
900 AsnGluValSerGluAspGlyValSerSerThrvAlaSpGluLeuSer 916
157 AATGAACTTCTGAAAGATGAGATCTCGAGTACTGAGCAACACTAG 206
916 pIleHstIleGluProGlyThrvAsnAspSerGlnHisSerLyCyAsp 933
207 CATCATATAGAGCTGGAACCATGATCTCAGCACTTAATGTGATG 256
933 AlAspLySerValGluProGluProPhePheHstLyValAlaHisSer 949
257 TAGATTAATCTGTCGAACCGGAAACCATTTTCCATAGAGTTCATTC 306
950 GlnHstAsnLeuValProGlnValGlnSerValGlnCySerProG 966
307 GAACCTGAACTTAGTCTCCTCAAGTTCATCGTTCAGTTCACCA 356
966 uGluSerPheAlaPheArgSerHisSerHisLeuProProLyAsn 983
357 AGAATCTTTCATTCGATCTCAGTCTGCACTTTCACCAAAAAA 406
983 snLyAsnSerLeuLeuIleGlyLeuSerThrglyLeuPheAspAla 999
407 ACAAAATTCCTTCTGCTGATTCGACTTTCACAGTCTGTTGATG 456
457 AACCCAAAGATGTTAAGACATGTTCACTCCAGATCTCAAAAGCT 506
1000 AsnProLyMetLeuArgThrvCySerLeuProAspLeuSerLyS 1016
1016 eArgThrvLeuMetAspValProThrvAlaGlyAspValAlaGln 1032
507 CAGAACCTTATGGCATGTTCCACCGTAGAGATGTTCTCAAGACAT 556
1033 LeuGluIleAspGluIleLyAsp...GluAsnIleLyGluGlu 1048
557 CTTCGATATAGTACCTTCGACAGATGAAACCTTACAGAGAACCT 606
1048 eArgSerGluAspIleValPheGluGluThrvAspThrvAspLeu 1064
607 CTGATTCGAAAGACATGTCCTGAAAGAACTGACAGATTTTACA 655
1065 LeuGlnAlaSerMetGluGlnLeuLeuArgGluGluProGly 1081
656 CTGCAGGCTCGATGGAACATGTTACTTAGGAGCAACACTGTG 705
1081 yrSerGluGlu 1085
706 ACAGTGAGAGAA 719

```

seq_name: gb_est2:BE881153

seq_documentation_block:

LOCUS BE881153

DEFINITION 601462024F1 NIH_MGC_69 Homo sapiens CDNA clone IMAGE:3894389 5',

ACCESSION BE881153

US-09-783-320-4

VERSION BE881153.1 GI:10329929

EST.

KEYWORDS

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@bts.fremail.nih.gov

Tissue Procurement: DCD/DIT/Gazdar

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM9683 row: p column: 06

High quality sequence stop: 712.

Location/Qualifiers

1. 847

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3894389"

/clone_lib="NIH_MGC_69"

/tissue_type="Large cell carcinoma, undifferentiated"

/lab_host="DH10B (phage-resistant)"

/note="Organ: Lung; Vector: pCMV-Sport6; Site: 1; NCBI:

Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.1 kb. Library constructed by Life

Technologies.

BASE COUNT 313 a 129 c 225 g 180 t

ORIGIN

alignment_scores:

Quality: 1075.00 Length: 298

Ratio: 4.151 Gaps: 11

Percent Similarity: 86.913 Percent Identity: 81.879

alignment_block:

US-09-783-320-4 x BE881153

Align seg 1/1 to: BE881153 from: 1 to: 847

475 ProGluArgGlnGlyGlnLeuAlaValGluArgAlaValGlnVal 491

3 CCAGAAAGGCAAAAGGCGACGCTGAGAAAGAGCTAAACAGTAGA 52

491 uGluPheLeuGlnArgLyArgGlnAlaMetGlnAsnLyAlaArgAla 508

53 AGAGTTCCTGACGAGAAAGGAGAGCTATGCAAGATTAAGCTGAGCG 102

508 lUcGlyHisMetValTyrLeuAlaArgLeuArgGlnIleArgLeuGln 524

103 AAGACATATGCTTATCTGCGCAAGACGTAGGCAAAATAGACTAGAT 152

525 PheAsnGluArgGlnGlnIleLyAlaLyLeuArgGlyGlyLyVal 541

153 TTCATATAGCGCCACAGATTAAGCCAAACTTCGTGTAAGAAAGAA 202

541 uAlaAsnHisSerGluGlyGlnGlySerGluGlnAlaAspMetArg 558

203 AGCTAATCATTTGAGAGGACAGAGAGAGAGAGAGAGAGAGAGAG 252

558 rGlySerIleGluSerLeuLyAlaHisAlaAspAlaArgAlaVal 574

253 GCAAAAGAAATCGAATCATAGAGGCGCATGCAAAATGCGCTGCTG 302

575 LeuLyGluGlnLeuGluArgLyArgLyGlyGlnAlaTyrGluArgGlu 591

Wed May 15 14:14:07 2002

us-09-783-320-4.1st

Page 10

1029 ArgGlnAspAsnLeuGluLeuLeuasp 1036
|||||
603 CGTCAAGACAACTCTGTAATAGAT 626
seq_name: gb_est1:A1378221

seq_documentation_block:

LOCUS A1378221 642 bp mRNA linear EST 18-MAR-1999
DEFINITION t65c01.x1 Soares_NhHMPU.S1 Homo sapiens cDNA clone IMAGE:2069472
3' similar to SW:NEK1_MOUSE P51954 SERINE/THREONINE-PROTEIN KINASE
NEK1; mRNA sequence.

ACCESSION A1378221
VERSION A1378221.1 GI:4188074
KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 642)

AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgap@remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 691 Std Error: 0.00
Seq primer: -400P from Glbco
High quality sequence stop: 437.

FEATURES
source
1..642
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2069472"
/clone_id="Soares_NhHMPU_S1"
/tissue_type="Soares_NhHMPU_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pTZ19-3D-Pac
(Pharmacia) with a modified polylinker. Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NDHM, pregnant uterus
NbHPU, and fetal heart NbH19M) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT 168 a 129 c 126 g 219 t
ORIGIN

alignment_scores:

Quality: 1023.00 Length: 215
Ratio: 4.895 Gaps: 2
Percent Similarity: 97.209 Percent Identity: 94.419

alignment_block:

US-09-783-320-4 x A1378221/rev

Align seg 1/1 to reverse of: A1378221 from: 1 to: 642

159 LeuAlaArgThrCysIleGlyThrProTyrTyrLeuSerProGluIleCys 175
|||||
642 CTGGCTTCGAACTTCAGGACCCCAATCTCTGTCACCTGGAATCTG 593
|||||
175 sgluAsnLysProTyrAsnAsnLysSerAspIleTrrPalaleuGlyCysV 192
|||||
592 TAAACCAACCCCTTACCAATATATAAGTGACATTGGGCTCTGGGGTGTG 543
|||||
192 alLeuTyrGluLeuGlyThrLeuLysHisAlaphGluAlaGlySerM 208

|||||
542 TCCCTTTATGAGCTGGGACCACTTACCATGGCTTTTGAGCTGACAGTA 493
|||||
208 eLysAsnLeuValLeuLysIleIleSerGlySerPheProProValSer 224
|||||
492 TGAACACCTGTGACTGAAATATATCTGAGATCTTTTCCACCTGTGCT 443
|||||
225 LeuHisTyrSerTyrAspLeuArgSerLeuValSerGlnLeuPheLysArg 241
|||||
442 TTGCATATTCCTATGATCTCCGACATTGGGTGTCTCAATATTAAAG 393
|||||
241 gAsnProArgAspArgProSerValAsnSerIleLeuGlyLysGlyPheI 258
|||||
392 AAATCTTAGGATGACCATCAGTCACTCCATATTGGAGAAAGGTTTAA 343
|||||
258 leaLysArgIleGluLysPheLeuSerProGlnLeuIleAlaGluL 274
|||||
342 TAGCCAAACGATTTGAAAGTTCTCTCTCCACGCTTATTCAGAAAGA 293
|||||
275 PheCysLeuLysThrPheSerLysPheGlySerGlnProIleProAlaL 291
|||||
292 TTTTGCTTAAACATTTTGCAGATTGGATGCACAGCTATACAGCTAA 243
|||||
291 sArgProAlaSerGlyGlnAsnSerIleSerValMetProAlaGlnLysI 308
|||||
242 AAGACCACTTCAGACAAACCTCGATTCTGTATGCTGCTCAGAAAA 193
|||||
308 leThrLysProAlaAlaLysTyrGlyIleProLeuAlaTyrLysLysTyr 324
|||||
192 TTACAAAGCCCTGCCGCTAAATATGGAATCTTACCATATATGAATAAT 143
|||||
325 GlyAspLysLysLeuHisGluLysLysProLeuGlnLysHisLysGlnAl 341
|||||
142 GGAGCTTAAATTAATTCACGAAAGAAACCACTGCAAAACATTAACAGGC 93
|||||
341 aHisGlnThrProGluLysArgValAsnThrGlyGluGluArgGlyLysI 358
|||||
92 CCATCAAACTCCAGAGAAAGAGTGAATCTGCGAAGAAAGCAAGGAAAA 43
|||||
358 leSerGluGluAlaAlaArgLysArgArgLeuGluIlePheIle 371
|||||
42 TATCTGAGGAAGCAGCAAGAAAGAAAGCGCTGGAATTAAT 2

seq_name: gb_est1:AM976435

seq_documentation_block:

LOCUS AM976435 654 bp mRNA linear EST 02-JUN-2000
DEFINITION EST38834 MAGE resequences, MAGN Homo sapiens cDNA, mRNA sequence.
ACCESSION AM976435
VERSION AM976435.1 GI:8167663
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 654)

AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gasparid,R., Gay,C., Holt
I.E., Saeed,A.I., Sharov,Y., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.

TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray

JOURNAL Unpublished (2000)

COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 360

Seq primer: Forward.
Location/Qualifiers

1..654
/organism="Homo sapiens"

FEATURES
source

```

/db_xref="taxon:9606"
/clone_lib="MAGE_ressequences_MAGN"
/note="Vector: pBluescriptSKm"
BASE COUNT      171 a      134 c      121 g      228 t
ORIGIN

alignment_scores:
    Quality: 1016.00      Length: 215
    Ratio: 4.861          Gaps: 0
    Percent Similarity: 97.209      Percent Identity: 95.349

alignment_block:
US-09-783-320-4 x AW976435/rev ..

Align seg 1/1 to reverse of: AW976435 from: 1 to: 654

157 ValGluLeuAlaArgThrCysIleGlyThrProTyrTyrLeuSerProG1 173
    ::::: ..::: ..::: ..::: ..::: ..::: ..::: ..:::
643 CTGAGAGGGGGCCGACCTGATGAGGGCCCACTACTGTCTCAGCTGA 594
173 uilecysgluasnlspProTyrAsnAsnLysSerAspIleTTPAlaLeuG 190
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
593 AATCTGTAAACAACT..TACATATATTAAG..GACATTGGGCTCTGG 546
190 lYcysValleuTyrGluLeuCysThrleuIshisAlaPheGluAlaGly 206
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
545 GGTGTGCTCTTATGAGCTGTGACACTTAACATCTTTTGAAGCTGGC 496
207 SerMetLysAsnLeuValleuLysIleIleSerGlySerPheProProVa 223
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
495 AGTATGAAAAAACCTGCTGACTGAAAGATATATCTGAGATCTTTCACCT 446
223 lSerleuIshisTyrSerTyrAspLeuArgSerLeuValSerGluLeuPheL 240
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
445 GTCTTGCAATATTCCTATGATCTCCGAGTTGGTGCTCAGCTATTTTA 396
240 yAspAsnProArgPArgPArgProSerValAsnSerIleLeuGluLysGly 256
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
395 AAGAAATCTTAGGATGACCATTCAGTCAATCTCATATTGGAGAAAGGT 346
257 PheIleAlaIysArgIleGluLysPheLeuSerProGluLeuIleAlaG1 273
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
345 TTATATGCCAAMCGCATTTGAAAGTTCTCTCTCCTCAGCTTATTCAGA 296
273 uGluPheCysLeuLysThrPheSerLysPheGlySerGluProIleProX 290
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
295 AGAATTTGTCTTAACACATTTTTCGAACTTTCGATCAGACCTATACAG 246
290 lAlaLysArgProAlaSerGlyLysAsnSerIleSerValMetProAlaGln 306
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
245 CTAAAGACCCAGCTTCAGGACAAACCTCGATTCTGTTATGCCCTGCTCAG 196
307 lysIleThrLysProAlaAlaLysTyrGlyIleProLeuAlaTyrLysLys 323
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
195 AAAATTACAAAGCCTCCCGCTAAATATGGAATACCTTTAGCATTAAGAA 146
323 sTyrGlyAspLysLysLeuHisGluLysLysProLeuGlnLysHisLysG 340
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
145 ATATGGAGATAAATAATATACAGAAAGAAACCACTGCACAAAACATTAAC 96
340 lAlaHisGlnThrProGluLysArgValAsnThrGlyGluGluLysArg 356
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
95 AGGCCCATCAAACTCCAGAGAAAGAGTGAATATCTGGAGAAAGAGGAG 46
357 lysIleSerGluGluAlaAlaArgLysArgLeuGluPheIle 371
    ||||| ..::: ..::: ..::: ..::: ..::: ..::: ..:::
45 AAAATATCTGAGAGACGACGACAAAGAGAGAGCTGGAAATTATTTATT 1

seq_name: gb_est2:BF541163
seq_documentation_block:
LOCUS      BF541163              845 bp      mRNA      linear      EST 11-DEC-2000

```

```

DEFINITION      602068802P1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:406765 5',
ACCESSION      mRNA sequence.
VERSION      BF541163
AUTHORS      BF541163.1 GI:11628544
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      NIH_MGC http://mhc.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL      Contact: Robert Strausberg, Ph.D.
COMMENT      Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: LICM907 row: d column: 02
High quality sequence stop: 672.
location/Qualifiers
    1..845
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="NIH_MGC_58"
    /tissue_type="hypertrophied"
    /lab_host="DH10B (T1 phage-resistant)"
    /note="Organ: Kidney; Vector: pDNR-LIB (Clontech); Site: 1:
    SfiI (ggcgctggcgc); Site: 2: SfiI (ggcgctggcgc);
    Double-stranded cDNA was prepared from cell line RNA. 5'
    and 3' adaptors were used in cloning as follows: 5'
    adaptor sequence: 5'-ATTCTAAGAGCGCGCGCGCATG-AT(30)BN-3'
    (where B = A, C, or G and N = A, C, G, or T). Average
    insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies
    contained inserts by PCR. This library was enriched for
    full-length clones and was constructed by Clontech
    Laboratories (Palo Alto CA)."
BASE COUNT      272 a      146 c      209 g      218 t
ORIGIN

alignment_scores:
    Quality: 1006.50      Length: 292
    Ratio: 4.142          Gaps: 9
    Percent Similarity: 83.219      Percent Identity: 78.082

alignment_block:
US-09-783-320-4 x BF541163 ..

Align seg 1/1 to: BF541163 from: 1 to: 845

913 GlnLeuSerAspIleHisIleGluProGlyThrAsnAspSerGlnHisSe 929
    ..::: ..::: ..::: ..::: ..::: ..::: ..::: ..:::
1 CAATTTAGTACATTCATATATAGAGCTGGAACCAATGATTCAGACACTC 50
929 rLysCysAspValAspLysSerValGlnProGluProPhePheHisLysV 946
    ..::: ..::: ..::: ..::: ..::: ..::: ..::: ..:::
51 TAAATGATCTAGATAGCTCTGTGCACCGGACCATTTTCCATAGG 100
946 aValHisSerGlnHisLeuAsnLeuValProGlnValGlnSerValGln 962
    ..::: ..::: ..::: ..::: ..::: ..::: ..::: ..:::
101 TGGTTCATCTGACACCTTGAACCTTAGTCCCTCAAGTTCAATCATTGAG 150
963 CysSerProGluGluSerPheAlaPheArgSerHisSerHisLeuProT 979
    ..::: ..::: ..::: ..::: ..::: ..::: ..::: ..:::
151 TGTTCACCAAGAAATCCCTGGCATTTGCATTCATCCCATTTACACAC 200

```

Wed May 15 14:14:07 2002

us-09-783-320-4.rst

Page 12

979 c.LysanllysaNlysaNserLeuLeuIIgLIleuSerThCllyeu 995
 201 AAAAAATATAAACAAAGATTCCTGCTGATGGAGCTTCAACGTGCTG 250
 996 PheaspAlaasnAnpPolysMeLleuAgtTrCySSerLeuProAsle 1012
 251 TTGTGATCAAAACCAACCAAGATGTTAAGAGATGTTCCTACTCCAGACT 300
 1012 uSerLyLeuPheAgtThrleuMetAspaAlProthrValGlyAspaVal 1029
 301 CTCAAAGCTGTGCAGAACCTTATGATGTTCACACCGTAGAGATGTC 350
 1029 rGlnAspaAsnleuGluIIleAspGlu.IIleYsaAspGluasnIIleYsgI 1045
 351 GTCAAGACATCTTGAAATAGATGAACATGAAAGATGAACATTAAGA 400
 1045 uGlyProSerAspSerGluAspIleValPheGluGluThrAspThrAspL 1062
 401 AGGACCTTGTGATTTCTGAAGACAT.GTGTTTAAAGAACTGACACAGTT 449
 1062 euGInGluLeuGlnAlaSerMetGluGlnLeuLeuAgtGluGlnProGly 1078
 450 TACAAAGAGCTGCAGCGCTCGATGAACAGTACTTAGGAACAAACCTGCT 499
 1079 GluGluTyrSerGluGluGluGluGluSerVal.LeuYs.AsnSerAspAl 1094
 500 GAACAATAACAGTGAAGAAAGAGTCAAGTCTTTGAAGAAACAGTATGTG 549
 1095 GluProThrAlaasnGlyThrAspValAlaIAspGluAspAsn.ProS 1111
 550 GAGCAACACTGCAAAATGGAGCAGATGGCAGATGAAATGACAAATCCCA 599
 1111 eSerGluSerAlaLeuAsnGluGluTyrPheIserAspAsnSerAspGly 1127
 600 GCASTGTAAAGGTGCTCTGACGAAGAAAGAGCTCAGATPACAGTATGCT 649
 1128 .GluIIAlaSerGluGluGluGluGluGluGluSerValPheasnHIsleGluG 1144
 650 GSAATAATGTGATGATGATGATGATGATGATGCTCTTTACCAATTTGAGA 698
 1144 IuleAgtGlyeHIsleuGluGluGluGluGluGluGluGluGluGluGlu 1160
 699 C....CGGAGATTCCTTCGGGGCAGACATGGGTTGGAAAAATTTTGGCGT 745
 1161 ValTyrGluYs.....IIeYsaI 1167
 746 GTTGAAGAAAAAAGAGCTCTGGGAAGAAGACATTTTGTGTTTGCCCA 795
 1167 aIIeHIsleuGluAspGluAspGluAsnIleGluIleCysSerIleValG 1184
 796 GCCTCATTTTGGGGGAG.....AAGGTTTTTA 821
 1184 IAsnIleLeuGly 1188
 822 CAAATTTATTTGGG 835

seq_name: gb_est1:AA419471

seq_documentation_block:

LOCUS AA419471 548 bp mRNA linear EST 16-OCT-1997
 DEFINITION zV030101.s1 Soares_NHMPu.S1 Homo sapiens cDNA clone IMAGE:752593.3'
 similar to SW:NR4.TRYB 808942 PUTATIVE SERINE/THEONINE-PROTEIN
 KINASE A ; mRNA sequence.
 ACCESSION AA419471
 VERSION AA419471.1 GI:2079189
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 548)
 Miller,T., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Kucaba,T., Lacy,M., Le,N., Lemmon,G., Maira,M., Martin,J., Moore,B.

TITLE
JOURNAL
COMMENT

Schellenberg, K., Steptoe, M., Tan, F., Thisting, B., White, Y., Wyllie,
T., Waterston, R., and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watsn.wustl.edu

This clone is available royalty-free through INMIL; contact the
IMAGE Consortium (infoimage@lmj.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: 41m3 fwd. Et from Amersham
High quality sequence spot: 430.

FEATURES
Source
Location/Qualifiers
1..548
/organism="Homo sapiens"
/db_xref="GDB:5975427"
/db_xref="taxon:9606"
/clone="IMAGE:752593"
/clone_lib=" Soares_NHMPU_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pTZ19-3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NHM, pregnant uterus
NBHP, and fetal heart NBHI19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT 146 a 109 c 101 g 192 t

ORIGIN

alignment_scores:
Quality: 931.00 Length: 182
Ratio: 5.115 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.451

Alignment_block:
US-09-783-320-4 x AA419471/reverse ..

Align seg 1/1 to reverse of: AA419471 from: 1 to: 548

```

190 GTCYSVALLEUTYTGULEUCYSTHTHEULYSNLSIALAPHENGLUALAGI 206
548 GGGTGTCTCCTTTATAGAGCTGTGTACACTTAACCACATCTTTGAAGCGGG 499
206 YSEMETLVSANLEUVALLEULYSILEELGEERGLYSERPHEROPROV 223
498 CAGTATGAAAAAACCCTGGTACTGAAAGATAATATCGATCTTTCCACCTG 449
223 ALSETLEINHISTYSERTYRASPDELAISERTLEUYVAISERGINLEUPHE 239
448 TGTCTTTGCATTATTCCTATGATCTCCGACGATTGGTGTCTCAGTTATT 399
240 LYSATGASNPROATGASPARTPROSERVALANSERTILENGULLYSGI 256
398 AAAAGAAATCCTTAGGATAGACCATCACTACACTCCATTATGGAGAAGG 349
256 YPHETLEIALYSATGLEULYSYPHELEUSERPROGINLEUILEALG 273
348 TTTTATGCCAACGACATGAAAAGTTCTCTCTCTCCTACGTTATTCAG 299
273 LUGLUPHCYSLEULTHRPHESERTLYSPHEGLYSERCLIPROLLEPRO 289
298 AAGAAATTTTGCTAANAACATTTTGCAAGTTGGATCACAGGCTATACCA 249

```

290 AAlATyArGPrOAlaSerGlyInaSerIleSerValMetProAlaI 306
 248 GCTAAAGACCACTTCAGCAAACTCGATTCTGTATGCTCTCA 199
 306 nlystIlePrlYsPrOAlaIaIaIyTyGlyIleProLeuAlaTyLysL 323
 198 GAAATATCAAGCCTGCGCTAAATATGAAATACCTTATGATATAGA 149
 323 ystTyGlyAspLysLysIleuHISGlyLysProLeuGlnLysSHLys 339
 148 AATATGAGATATAAAATACCGAAAGAACACACGCAAAACATAA 99
 340 GlnAlaHISGlnThrProGlyLysArgValAsnThrGlyGluGluArg 356
 98 CAGGCCATCAACCTCCAGAGAGAGTGAATCTGAGAGAGAAAGCAG 49
 356 GlysIleSerGluGlnAlaIaIaIaIaIaIaIaIaIaIaIaIaI 371
 48 GAAATATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTT 3

seq_name: gb_est1:AA907081

seq_documentation_block:

LOCUS AA907081 477 bp mRNA linear EST 26-AUG-1998
 DEFINITION O103a06.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1522354.3,
 similar to SW:NEK1_MOUSE P51954 SERINE/THREONINE-PROTEIN KINASE
 NEK1 ; mRNA sequence.

ACCESSION

AA907081 GI:3042541

VERSION

AA907081.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 477)

NCI-CGAP

http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapsr@emil.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.S.E. website: www-bio.llnl.gov/bdrp/image/image.html

Insert Length: 461 Std Error: 0.00

Seq primer: -40ml3 fwd. RT from Amersham

High quality sequence stop: 373.

Location/Qualifiers

1..477

/db_xref="taxon:9606"

/clone_image="1522354"

/clone_id="NCI_CGAP_Lu5"

/tissue_type="carcinoid"

/lab_host="DH10B"

/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a

modified polylinker; 1st strand cDNA was prepared from

neuroendocrine lung carcinoid, and was then primed with a

Not I - oligo(dT) primer. Double-stranded cDNA was ligated

to Eco RI adaptors (Pharmacia), digested with Not I and

cloned into the Not I and Eco RI sites of the modified

pT73 vector. Library is normalized. Library was

constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

124 a 90 c 89 g 174 t

ORIGIN

alignment_scores:
 Quality: 813.00 Length: 159
 Ratio: 5.113 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-783-320-4 x AA907081/rev ..

Align seg 1/1 to reverse of: AA907081 from: 1 to: 477

213 leuyllelleSerGlySerPheProProValSerLeuHISySerTy 229
 477 CTGAAGATATATCTCGATCTTTCCACCTGTGCTTGCATATTCCTA 428
 229 raspleuArgSerleuValSerGlnLeuPheLysArgAsnProArgSPA 246
 427 TGAATCTCGAGATTGTGTCTCATTTTAAAGAAATCTAGGAGATA 378
 246 rpfroSerValAsnSerlelleuGlySGlyPheIleAlaLysArgyle 262
 377 GACCATCAGTCAACTCATATTGAGAAAGTTTATAGCCAAAGCAT 328
 263 GluLysPheLeuSerProGlnleuIleAlaGluGluPheCysIleLysTh 279
 327 GAAAGATTCTCTCTCTCTCACTTATGCGAGAGATTTTCTTAAAC 278
 279 rPheSerLysPheGlySerGlnProIleProAlaLysArgProAlaSerG 296
 277 ATTTTCAGAGTTTGATCACAGCTTACCACCTTAAAGACCACTTCAG 228
 296 lYgInaAsnSerIleSerValMetProAlaGlnLysIleThrLysProAla 312
 227 GACAAACTCGATTCTGTTATCCCTGCTCAGAAATTAACAAAGGCTCC 178
 313 AAlATyArGlyIleProLeuAlaIaIaIyTyGlyAspLysLysLe 329
 177 GCTAATATGGAATACCTTATGATATAGAAATATGAGATTAATAATT 128
 329 hISGlnLysLysProLeuGlnLysHISLysGlnAlaHISGlnThrProG 346
 127 ACACGAAAGAAACACACGCAAAACATTAACAGGCCCTCAACTCCAG 78
 346 lAlaArgValAsnThrGlyGluGluArgArgLysIleSerGluGluAla 362
 77 AGAAGAGACTGATATCTGAGAGAAAGAGAAATATCTAGGAGACA 28
 363 AAlATyArGlyArgArgLeuGluPheIle 371
 27 GCAGAGAGAGAGAGAGCTGGAATTTATT 1

seq_name: gb_est1:AL036359

seq_documentation_block:

LOCUS AL036359 544 bp mRNA linear EST 29-FEB-2000
 DEFINITION DKFZP564C153.r1 564 (synonym: hfbz2) Homo sapiens cDNA clone
 DKFZP564C153.5', mRNA sequence.

ACCESSION

AL036359

VERSION

AL036359.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 544)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

Wambuit, R., Haubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

EST (Wambuit, et al.)

Unpublished (1999)

CONTACT: Wambuit R

MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert

clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

THIS PAGE BLANK (USPTO)


```

117 shisValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsn 134
118 |||||||
119 |||||||
120 ACATGATACATGATGAAATTTCTTCAGACATTAATATCAGACACA 457
121 |||||||
122 |||||||
123 LePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
124 |||||||
125 |||||||
126 TATTTTAACTAAAGATGAGACACTACACTGGAGATTTTGGAAATTCG 507
127 |||||||
128 |||||||
129 ArgValIleAsnSerThrValGlnLeuAlaArgPheCysIleGlyThrP 167
130 |||||||
131 |||||||
132 ACAGTCTTAAATAGACTGATGAGCTGGCTGAACTTGCAATGAGGACCC 557
133 |||||||
134 |||||||
135 CTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 184
136 |||||||
137 |||||||
138 ATACTACTGTCACCTGAAATCTGTAACAAACCTTACAAATTAATAAA 607
139 |||||||
140 |||||||
141 eAspIleThrPalalLeuGlyCysValLeuTyrgIleuGlyCysThrLeu 200
142 |||||||
143 |||||||
144 GTGACATTTGGGCTCTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 657
145 |||||||
146 |||||||
147 HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleIle 217
148 |||||||
149 |||||||
150 CATGCTTTTGAAGCTGGCAGTATGAAACCTGCTACTGAGAAATATATC 707
151 |||||||
152 |||||||
153 fGlySerPheProProValSerLeuHisTygSerTygAspLeuArgSer 234
154 |||||||
155 |||||||
156 TGGATCTTTTCCACCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 757
157 |||||||
158 |||||||
159 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAs 250
160 |||||||
161 |||||||
162 TGGTGTCTCAGTATTTAAAGAAATCTAGGGATACCATCATGCTAAC 807
163 |||||||
164 |||||||
165 SerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPheLeu 267
166 |||||||
167 |||||||
168 TCCATATTTGGAGAAAGTTTATAGCCAAAGCAGTGAAGATTTCTCTC 857
169 |||||||
170 |||||||
171 fProGlnLeuIleAlaGluGluPheCysLeuLysThrPheSerLysPhe 284
172 |||||||
173 |||||||
174 TCCCTCAGCTTATTCAGAAAGATTTTGTAAACAAATTTTCGAAAGTT 907
175 |||||||
176 |||||||
177 ySerGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
178 |||||||
179 |||||||
180 GATCACACCTATACACCTTAAAGACAGCTTCAGACAAACCTGCATT 957
181 |||||||
182 |||||||
183 SerValMetProAlaGlnLysIleThrLysProAlaIleAlaLysTyg 317
184 |||||||
185 |||||||
186 TCTGTTATGCGCTGCAGAAATTTACAAAGCCCTCCGCTAAATATGAA 1007
187 |||||||
188 |||||||
189 eProLeuAlaTyrgLysTyrgLysAspLysLysLeuHisGlnLysLys 334
190 |||||||
191 |||||||
192 ACCTTAGATATTAAGAAATATGGAGATTAATAATTTACAGAAAGAAC 1057
193 |||||||
194 |||||||
195 fGlnLeuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgVal 350
196 |||||||
197 |||||||
198 CACTGCAAAACATTAACAGCCCATCAACTCCAGAGAAAGAGCTGAAT 1107
199 |||||||
200 |||||||
201 ThrGlyGluGluArgArgLysIleSerGlnGluAlaAlaArgLysArg 367
202 |||||||
203 |||||||
204 ACTGGAAGAAAGAGAGAAATATCTGAGGACACAGAAAGAGAGAAAG 1157
205 |||||||
206 |||||||
207 gLeuGlnPheIleGluLysGluLysGlnLysAspGlnIleIleSerL 384
208 |||||||
209 |||||||
210 GCTGGAATTTTATGAAAAAGATAGGAACCGTAGATCAGATTAATTA 1207
211 |||||||
212 |||||||
213 euMetLysAlaGluGlnMetLysArgGlnGluLysGluArgLeuGluArg 400
214 |||||||
215 |||||||
216 TAAATGAGGCTGAACAATGAAAGGCAAGCAAGAAAGTTGGAGAGA 1257
217 |||||||
218 |||||||
219 IleAsnArgAlaArgGluGlnGlyThrPheArgValIleuSerAlaGly 417
220 |||||||
221 |||||||
222 ATTAATAGGCGCAGGAAACAAGATGAGAAATGTCTTAAGTCTGTG 1307

```

```

417 ySerGlyGluValLysAlaProPheLeuGlySerGlyGly.....Thr 432
418 |||||||
419 |||||||
420 AAGGTGAGAGTAAAGTAGGATTTTATATACCAATATGTTTACTACCA 1357
421 |||||||
422 |||||||
423 LeAlaProSerSer 436
424 |||||||
425 |||||||
426 TTTTCCCTCCACAT 1371
427 |||||||
428 |||||||
429 seq_name: /cgn2_6/ptodata/1/ina/6B_COMB.seq:us-09-420-915-15
430
431 seq_documentation_block:
432 ? Sequence 15 Application US/09420915
433 ? Patient No 6264947
434 ? GENERAL INFORMATION:
435 ? APPLICANT: Bandman, Olga
436 ? APPLICANT: Tang, Y. Tom
437 ? APPLICANT: Hillman, Jennifer L.
438 ? APPLICANT: Yue, Henry
439 ? APPLICANT: Guebler, Karl J.
440 ? APPLICANT: Corley, Neil C.
441 ? APPLICANT: Gorgone, Gina
442 ? APPLICANT: Azimzal, Yalda
443 ? APPLICANT: Lu, Alina
444 ? TITLE OF INVENTION: Protein Kinase Homologs
445 ? FILE REFERENCE: PF-0614 US
446 ? CURRENT APPLICATION NUMBER: US/09/420,915
447 ? EARLIER FILING DATE: 1999-10-20
448 ? EARLIER APPLICATION NUMBER: US 09/173,581
449 ? NUMBER OF SEQ ID NOS: 18
450 ? SOFTWARE: PERL Program
451 ? SEQ ID NO 15
452 ? LENGTH: 1846
453 ? TYPE: DNA
454 ? ORGANISM: Homo sapiens
455 ? FEATURE: -
456 ? OTHER INFORMATION: 1567782
457 ? US-09-420-915-15
458
459 alignment_scores:
460 ? Quality: 1959.00 Length: 438
461 ? Ratio: 4.922 Gaps: 2
462 ? Percent Similarity: 90.868 Percent Identity: 89.269
463
464 alignment_block:
465 ? US-09-783-320-4 x US-09-420-915-15 ..
466
467 Align seg 1/1 to: US-09-420-915-15 from: 1 to: 1846
468
469 1 MetGluLysTyrgValArgLeuGlnLysIleGlyGluGlySerPheGly 17
470 |||||||
471 |||||||
472 154 ATGAGAGAGATGTGATGACTACAGAAAGATTGGAGAAAGTTCTTTG 203
473 |||||||
474 |||||||
475 17 sAlaIleLeuValLysSerThrGlnAspGlyArgGlnTyrgValIle 34
476 |||||||
477 |||||||
478 204 AGCCATTCCTGTTAATCTACAGAGATGGCAGACGATGTGTTCAAG 253
479 |||||||
480 |||||||
481 34 IuIleAsnIleSerArgMetSerSerLysGluArgGluGluSerArg 50
482 |||||||
483 |||||||
484 254 AAATTAACATCTCAAGAAATGCTCAGTAAAGAAAGAGAAAGAAAT 303
485 |||||||
486 |||||||
487 51 GluValAlaValLeuAlaAsnMetLysHisProAsnIleValGlnTy 67
488 |||||||
489 |||||||
490 304 GAACTTCAGTATTTGGCAACATGAGACATCCAAATATTGTCAGTAA 353
491 |||||||
492 |||||||
493 67 gGluSerPheGluGlnLysnGlySerLeuTyrgIleValMetAspTy 84
494 |||||||
495 |||||||
496 354 AGATCATTTGAA..... 366
497 |||||||
498 |||||||
499 84 IuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeu 100
500 |||||||
501 |||||||
502 366 ..... 366

```

```

101 GINGLUAAPGLNILEUAAPTRPHEVALGNIIECYSEUVALALEU 117
117 SHSVAHISAPARGLYSILEUHIASARGASPILELYSSEGLNANI 134
408 ACATGACATGATGAAAAATTCCTCAGACATTAATTCACAGACA 457
134 IEPHEUHLRYSPGLYTHRVALIGNLEUGLYASPPHEGLYILEA 150
458 TATTTTAACTAAAGATGACAGACACTTGAGATTTTGGAAATGCT 507
151 ARGVALLEUANSERTHVALIGNLEUVALARGTHCYSTIEGLYTHPR 167
508 AGAGTCTTAATAGTACTGATGAGCTGCTGCACTGCTGCAATGAGGACCC 557
167 CTYTYRLEUSERPROGILIECYSGIUNASLYSPROLYTASPNLYS 184
558 ATACTACTTGTCTACCTGAAATCTGTGAATAACAACTTACATAATATAA 607
184 ERASPILETTPALALEUGLYCYSVALLLEUYRGILEUCYSTHLEULYS 200
608 CTGACATTTGGCTCTGGGGTGTCTTATGAGCTGTGACACTTAA 657
201 HISLAPHEGLUAGLYSERMELYSASNLEUVALLEULYSILEIIESE 217
658 CATGCTTTGAGCTGCGAGTATGAAAAACCTGTGACTGAAGATAATATC 707
217 TGLYSERPHROPROVALSERLEUHIISTYISERTYRASPLEUARGSERL 234
708 TGGATCTTTTCCACCTGTCTTTCATATTCATATGATCTCCCACTT 757
234 EUVALSERGINLEUPHELYSARGASNPROARGASPARPROSERVALASN 250
758 TGGTGTCTCAGATTATTAAGAAATCTAGAGATGACATCACTGAC 807
251 SERILEUGLIULYSGLYPHEILEALYSARGILEGLYSPHELESE 267
808 TCCATATTTGAGAAAGTTTATAGCCAAACGCAATGAAAGTTTCTCTC 857
267 PROGLINEULEALAGLUPHECYSEULEYSTRHPHSELYSPHSG 284
858 TCCCTAGCTATTCACAGAAATTTTCTTAAAAACATTTTGAAGTTTG 907
284 IYSEGLNPROILEPROALALYSARGPROALASERGLYGNASNSERTILE 300
908 GATCACAGCTATACAGCTAAAGACAGCTTCAGACAAAACTGAT 957
301 SERVALMETPROALAGLILYSILETHLYSPROALALALYSTYGLY 317
958 TCTGTATGCCCTGCTCAGAAAAATTAACAAGCTGCCCTTAATATGAAAT 1007
317 EPROLEUALATYRILYSLYSTYRGLYASPLYSLEUHIISGLULYSERP 334
1008 ACCTTTAGCATTAAGAAATATGAGATTAATAATTTACAGAAAGAAAC 1057
334 FOLEUGLILYSILYSGLINAHLSGLNTHRPROGLULYSARGVALASN 350
1058 CACTGCAGAAAACATAACAGGCCCTCAAACTCCAGAGAGAGATGAAT 1107
351 THGLYGLUGLUARGARGLYSILESERGLINGLUALAARGLYSARGAR 367
1108 ACTGGAAGAAAGAGAGGAAATATCTGAGAGACAGACAGAAAGAGAAAG 1157
367 GLEUGLUPHEILEGLULYSGLULYSGLINLYSPOLINLEILESERL 384
1158 GCTGGAAATTTATGAAAAAGATTAAGAAACGATGAGATCAGATTTTATGTT 1207
384 EUWELYSALAGLUGLIMELYSARGGLINGLULYSGLIARGLEUGLARG 400
1208 TAATGAAGCTGAACAATGAAGAAAGCAAGCAAGAAAGTTGGAAGA 1257

```

```

401 ILEASARGALARGLUGLINGLYTRPARGASNVALLEUSERALAGLY 417
1258 ATRAAATAGGCCGAGCAGACAGATGAGAAATGCTCTAGTCTGTGG 1307
417 YSERGLYGLUVALYSALAPROPHLEUGLYSERGLYGLY.....THI 432
1308 NAGTGTGAAGTAAAGTAGGCAATTTATACATATAGTTATACACCA 1357
432 LEALAPROSER 436
1358 TTTTCCCTCCACT 1371

seq_name: /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-08-870-529-1
seq_documentation_block:
; Sequence 1, Application US/08870529
; Patent No. 6080557
; GENERAL INFORMATION:
; APPLICANT: Sims, John E.
; APPLICANT: Virca, G. Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: IL-1/TNF- $\alpha$ -ACTIVATED KINASE (ITAK),
; TITLE OF INVENTION: AND METHODS OF MAKING AND USING THE SAME
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870.529
; FILING DATE: 06-JUN-1997
; CLASSIFICATION: A35
; ATTORNEY/AGENT INFORMATION:
; NAME: McMaisters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.418
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2940 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2937
; US-08-870-529-1

alignment_scores:
Quality: 600.00 Length: 392
Ratio: 2.353 Gaps: 8
Percent Similarity: 65.051 Percent Identity: 33.673

alignment_block:
US-09-783-320-4 x US-08-870-529-1 ..
Align seg 1/1 to: US-08-870-529-1 from: 1 to: 2940

4 TYVALARGLEUGLILYSILIEGLYGLUGLISERPHGLYYSALALE 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
154 TACATCCCATTCGCGTCTGTGGCGCGCGGCTTGGGGAAGCCAGCT 203

```



```

296 .....G 296
1158 AGCGAAGCTATGTTGGGCTGGTGGAAATCCACCCCGAGAAACTGGA 1207
296 ..
296 1yqlnlsnserlleserValmetProIaglnLysIleThrLysProIa 312
1208 TGTATTCAGAGAGCTGTATAGTCCCGGAGGTGTGTG..... 1245
313 AlAlaLysTyrLysLysProLeuAlaTyrLysLysTyrGlyAspLysL 329
1246 .....CAGGAATACCACCTTTCGTGTGTGTGCACAGTGGAGAGCAATG 1289
329 euHISGulLysProLeuGlnLysHisLysGlnAlaHisGlnTrpPro 345
1290 TACAC.....TTGGCGAACAlTCAGCAAGGAGGCACCTAACTCCA 1327
346 .....GulLysATGValIleThrGlyLugIleAr 355
1328 TGTGCACCTGGCCCATGGAGACAAAGCTCTATGCACAGCCAAAGCATG 1377
355 gArGlyIleSerGluGluAla 362
1378 TGGAAATGTTCCAGAGCCAAAGC 1399
seq_name: /cgn2_6/ptcdelta/1/lna/6B_COMB.seq:US-09-393-569-1

```

```

seq-documentation block:
Sequence 1, Application US/09393569
Patent No. 6279979
GENERAL INFORMATION:
APPLICANT: BINGHAM, SHARON
APPLICANT: CASE, PAMIRK
APPLICANT: LAMSON, SALLY NEALE
APPLICANT: NENTON, RICHARD ANTHONY
APPLICANT: PIERCE, VALERIE
APPLICANT: RAUSCH, OLIVER LARS
APPLICANT: REITH, ALASTAIR DAVID
APPLICANT: SANGER, GARETH JOHN
TITLE OF INVENTION: NEW USE
FILE REFERENCE: P32201261
CURRENT APPLICATION NUMBER: US/09/393,569
CURRENT FILING DATE: 1999-09-10
EARLIER APPLICATION NUMBER: GB 9907261.3
EARLIER FILING DATE: 1999-03-29
EARLIER APPLICATION NUMBER: GB 9819779.1
EARLIER FILING DATE: 1998-09-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: SEQ ID for Windows Version 3.0
SEQ ID NO: 1
LENGTH: 4181
TYPE: DNA
ORGANISM: HOMO SAPIENS
US-09-393-569-1

alignment_scores:
Quality: 493.00 Length: 1242
Ratio: 0.824 Gaps: 61
Percent Similarity: 48.148 Percent Identity: 22.866

alignment_block:
US-09-783-320-4 x US-09-393-569-1 ..

Align seq 1/1 to: US-09-393-569-1 from: 1 to: 4181

7 leuclnlyslleclgluclyserpheglylysajalleuevallysse 23
:||||:||||:||||:||||:||||:||||:||||:||||:||||:
180 GTGCAGACTTGTTCGAATGCAACATVCGGACAGATTATAGCGTCGCA 22
23 rthclaspcllyarlglntrvallelysclyllesnllserarqm 40
:||||:||||:||||:||||:||||:||||:||||:||||:

```

230 TGTCAAAAGGGGCAAGCTTTCAGCCATCAAGGTTATGATGTC..... 272
40 eTserSerLysGluArgGluGluSerArgGluValAlaValLeuAla 56
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
273 ..ACGGGGATGAAAGAGAGAAATCAACAAGAAATTAACATGTTAGAG 320
57 AsnMetLys...HisProAsnIleValGlnTyrArgGluSerPheGluG 72
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
321 AAATATTCTCATCACCGGAATATTGTACATACATGCTGCTTTTTCAA 370
72 u.....AsnGlySerLeuTyrIleValMetAspArg 83
:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
371 AAAGAACCCACCCAGCATGATGACAACTTGTGTTGTCATGAGCTTT 420
83 ySGlGluGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeu 99
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
421 GTGGTGGCTGGCTGTGTACCCAGCTGATCAAGAACCAAAAGGTAACAG 470
100 PheGlnGluAspGlnIleLeuAspTyrPheValGlnIleCys..... 113
471 TTGAAGAG.....GAGTGGATTGTCATACATCTGCAGGGAAAT 508
114 ...LeuAlaLeuLysHisValHisAspArgLysIleLeuHisArgAsp 129
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
509 CTTACGGGGGCTGAGTCACTGCACACGATMAAGTATGATCATGAGATA 558
129 IeLysSerGlnAsnIlePheLeuThrLysAspGlyThrValGlnLeuGly 145
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
559 TTAAGGGCAAAATGTCTTCTGACTGAATAAGCAAGTAACATAAGT 608
146 AspPheGlyIleAlaArgValLeuAsnSerThrValGluLeuAlaArgH 162
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
609 GACTTGGAGTCAGTCCAGCTTGCATGCAACAGTGGGAGAGAGATAC 658
162 fCysIleLeuThrProTyrTyrLeuSerProGluIle.....CysGlu 177
||||||:::|||||:::|||||:::|||||:::|||||:::|||||
659 TTTCATGGAACCTCCATCTGATGGCAACCAAGATGATGCTGCTGATG 708
177 snLysPro.....TyrAsnAsnLysSerAspIleTrrAlaLeuGly 190
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
709 AAAACCCAGATGCCACATATGATTTCAAGAGTCACTTGTGCTTTGGG 758
191 CysValLeuTyrGluLeuCysThrLeuLysHisAlaPheGluAlaGly 207
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
759 ATCACCSCCATTTGAATGGCA..... 779
207 fMetLysAsnLeuValLeuLysIleIleSerGlySerPheProValS 224
780GAAAGTGGTCCCCCTGCTGTG 801
224 eLeuHisTyrSerGlyAspLeuArgSerLeuValSerGlnLeuPheLys 240
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
802 ACATGCAC.....CCCATGAGAGCTCTCTTC.....CTCATCCSC 836
241 ArgAsnProArgAspArgProSerValAsnSerIleLeuGlnLysGlyH 257
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
837 CGAAATCCAGCGCTGG.....CTGAAGCTTAAGAA 868
257 eIleAlaLysArgIleGluLysPheLeuSerProGlnLeuIleAlaGlu 274
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
869 GTGGCAAAATAATCCAGTCAATTATTGAGAGCTGCTGTAAGAAAC 918
274 IuPheCysLeuLeuThrPheSerLysPheGlySerGlnProIleProAla 290
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
919 AC.....AGC 923
291 LysArgProAlaSerGlyGln.....AsnSerIleSerValMet 303
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
924 CAGCGACGACGACAGAAATGATGAGACATCCATTTATGACAGACCA 973
303 fProAlaGlnLysIleThrLysProAlaAlaLysTyrGlyIleProLeu 320
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
974 ACCTAATAGAGGACAGGTCCGATTCACATCAAGGACCAATATTGATGA 1023

320 IatTyrLysLysTyrGluAspLysLysLeuHisGluLysLysProLeuGln 336
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1024 CAAAGAAAGAGACGAGGA.....GAAAAGATGACAGACAGAG 1058
337 LysHisLysGlnAlaHisGlnThrProGluLysArgValAsnThrGlyG 353
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1059 TATGATGACAGCTGAGATGAGGAAGAGAGAGAGATGACTCGAGGA 1108
353 uGlu.....ArgArg.... 356
||||||
1109 CCCCAGCTCCATCTGATCTGCAAGGGAGTGCAGCCTGGAGAGAT 1158
357LysIleSerGlnGlnAlaAlaArgLys 365
1159 TTCTGAGCTGCAGCTGGCCCAACAGAGCCTTCTGAGGCTCTAGGAG 1208
366 ArgArgLeuGlnIlePheIleGluLysLysGln...LysAspGlnI 381
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1209 CAGAGCTGGAGCGACGACGCGGAGATGAGAGCAACAGCCGACACT 1258
381 eIleSerLeuMetLysAlaGluGlnMetLysArgGlnLysGluArgL 398
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1259 GCTG.....GCCGAGCGTCAAGGCGCATCGAGGACGAGAAAG 1296
398 euGluArgIleAsnArgAlaArgGluGlnGlyTrrArgValLeuSer 414
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1297 AGCAGAGG...CGCGGCTGGAGGAGCAACAAAGCGA..... 1331
415 AlAGLyGlySerGlyLysValLysAlaProPheLeuGlySerGlyTh 431
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1332GAGAAAGACTGCCGAAGCAG..... 1352
431 rIleAlaProSerSerPheSerSerArgGlyGlnTyrGlnHisTyrHis 448
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1353CAGAGAGGAGACAGCGCGGCACTATGAG. 1382
448 IatIlePheAspGlnMetGlnGlnAlaArgAlaGluAspAsnLys 464
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1383GAGCAGATGCGCGGAGAGAG...GAGAGAGCGGCTCGGAG 1421
465 TrpLysArgGluIleTyrGluArgGlyLeuProGluArgGlnLys.... 479
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1422 CATGAAACAGGAATGATCATGACGCGCAGTTGAGAGAGACAGACACT 1471
480GlyIleLeuAlaValGluArgAla..... 487
1472 AGAGATCTGCACAGCAGACTACTGCATGACACACTCTACTTGAAAT 1521
488LysGlnValGluGluPheLeuGlnArgLysAlaGlnAlaMet 501
1522 ATGAAGCCAAACAAATTTGGAAGAA.....CAGACACAAACAGAAAGCTG 1565
502 GlnAsnLysAlaArgAlaGlu...GlyHisMetValTyrLeuAlaArgLe 517
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1566 CAGAGCGACCTAAAGCAAGAAAGAGACTACTAGTTCCCTTCACAGCA 1615
517 uArgGlnIleArgLeuGlnAsnPheAsnGluArgGlnIleLeuAlaL 534
||||||:::|||||:::|||||:::|||||:::|||||:::|||||
1616 GCGGCAAG.....GAGCAGAGCGCTGTG..... 1637
534 yLysLeuArgGlyGluLysGluAlaAsnHis...SerGlnGlyGln 549
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1638GAGAGAAAGCCCTGATCAATTACAAAGAGAAATGAGT 1676
550 GlySerGlnGluLysAlaAspMetArgLysLysIleGlySerLeuLysAl 566
|||||
1677 CCTAGTGA..... 1685
566 aHisAlaAsnAlaArgAlaValLeuLysGlnGlnLeuGlnLysAla 583
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1686AAGCCAGATGGGSCAAAGAGAGTAAAGAAAGCGTGA 1722

583 rglYsgluAlaTyrGluArgGluTyrLysValTrpGluGlnHisLeuVal 599
|||
1723 GG.....CTCAACGGGAAAGTTCCCTCCATGCTCCACAAAGTT 1763
|||
600 AlaLysGlyValLysSerSerProLeuGlnHisG1 616
|||
1764 GCCACAGGATATCTGACCCCAACCTGCCCAAGGTGAGTCTTCAG 1813
|||
616 urhnglyGly.....SerProSerL 623
|||
1814 CATTAGTGGAGTTCCAGCTGCTGCAACCCCCCTGCTCAGACAGTCG 1863
|||
623 ysgIngluMetArgSerValHisSerValHisLeuLysGluVal 639
|||
1864 ATCCCGAGATCCCATCTGGTAGCTGTAATAATCCAG..... 1901
|||
640 G1ValAlaSerSerLeuThrAspThrArgGluTrpSerGluGluMetG1 656
|||
1902GGACCTGCTTGACCGCTCCAGTCCAGTCCAGACAGACCCAC 1945
|||
656 nlygThrAspAsnAlaHisSerSerLysArgGluLeuArgArgLeuA 673
|||
1946 AAAG.....GGCTCTCTGGGTTTCAGGAGGCTCTG..... 1976
|||
673 snGluAsnLeuLysAlaGlnGluAspGluLysGlyMetGlnAsnLeuSer 689
|||
1977AACGTGACCTCCACCGCTGAGATGCCACGCGCAAGATCAGAT 2021
|||
690 AspThrPheGluHisLeuAsnValHisGluAspAlaLysGlnHisGluLysG1 706
|||
2022 CCCACCTCGGAA.....AATCTCTCTCTCCCCAC 2050
|||
706 nlygSerValSerSerAspArgLysTrpGluAlaGlyGlnLeuVal 723
|||
2051 TCGCATGTAAGTTGACCGACGACCTCTGTATACGACAGAGAGAGACA 2100
|||
723 allLePro...LeuAspGluLeuThrLeuAspThrSerPheSerThrThr 738
|||
2101 TTCACCAAGAGTCCCTCAAGAACACTTATATCCACAGATTAGCC 2150
|||
739 GlnArgHisThrValGlyGluValHisLysLeuGluPro.....AsnG1 753
|||
2151 AGAAGAAATTCCTCGGATGCTGATGCTGTGGACCCAGACTAGATC 2200
|||
753 ySerProArgArgAlaTrpGlyLysSerPro.....ThrAspS 766
|||
2201 TCACCCATCAGAGCA.....AGCACCTGATCTCCGAGAACTGAGC 2244
|||
766 erValLeuLysIleLeuGluAlaGluLeuGln..... 777
|||
2245 CCAATCTTG.....GAGAGCCCTTGCAGAGGACGACGAGTGC 2282
|||
777 777
|||
2283 AGTTCTCCAGCTCAGCACCCCTAGCTCCAGCCCACTCCCAAGAG 2332
|||
778LeuGlnThrGluLeuLeuGlnAsnThrThrIleArgS 790
|||
2333 CTCACCACTGATCACAAGCAGAGATCCAGTGAACGACAGAGTTCGAG 2382
|||
790 erGluIleSerProGluGlyGluLysTyrLysProLeuIleThrGlyGlu 806
|||
2383 CCAACAGTAACTCAGAAAGATCACCTGCTTCC..... 2417
|||
807 LysLysValGlnCysIleSerHisGlu.....IleAsnProSerAl 820
|||
2418CATGAGCTTGCACAGGTGTAACCCGAGAA 2446
|||
820 allLeValAspSerProValGluThrLysSerProGluPheSerGluAlaS 837
|||
2447 ATCCAGGACATTCACCGGCCAGTGCACAGCTACTCAAAAAAGCTA 2496
|||
837 erProGlnMetSerLeuLysLeuGlnGlyAsnLeuGlnGluProAsp 853

2497 TAGATGAGATCTGACGACGATTACCCAAAGAACTAAGGAA..... 2537
|||
854 LeuGlnThrGluIleLeuGlnGluProSerGlyThrAsnLysAspGluSe 870
|||
2538 CTCGGATTGAAACAAACCCCAATGAAAGAGTGAATGATCTACTC 2587
|||
870 rLeuProCysThrIleThrAspValTrpLeSerGlnLysGluThrL 887
|||
2588 CTCCTCC.....AGTGAAGATCAGAAAGTA 2613
|||
887 ysgLutrhGlnSerAlaAspArgIleThrIleGlnGluAsnGluValSer 903
|||
2614 GCGAG.....GAAGAGAG 2627
|||
904 GlnAspGlyValSerSerThrValAsp.....GlnLeuSerAspI1 917
|||
2628 GAGATGTGAGAGACGAGACCATGATGGACAGATGCTGACGAGCAT 2677
|||
917 eHisIle.....GluProGlyThrAsnAspSerGlnHis 929
|||
2678 ACCCAGACTGATACCAACAGGAGCTCCAGGACGACAGAGATCAATG 2727
|||
929 erLysCysAspValAspLysSerValGlnProGluProPheHisLys 945
|||
2728 TCGCAATGTGGGAGCCATGCGCTGCGAGACTCT..... 2762
|||
946 ValValHisSerGlnHisLeuAsnLeuValProGlnValGlnSerValG1 962
|||
2763CATCGGACAGTTTC.....AGCGCGAG 2785
|||
962 nCysSerProGluGlnSerPheAlaPheArgSerHisSerHisLeuProP 979
|||
2786 TATTTCAGAGAGAGAACCTTGATGATTAGAGAGACTCT..... 2825
|||
979 rOlyAsnLysLysLysAsnSerLeuLeuIleGlyLeuSerThrGlyLeu 995
|||
2826GGAGAGAGAGGATCT.....GCCGAC 2849
|||
996 PheAspAlaAsnAsnProLysMetLeuArgThrCysSerLeuProAspLe 1012
|||
2850 AGTGACAGCAATGGC.....TTTCTGCGCAATCAACCTCCCTGACCT 2893
|||
1012 uSerLysLeuPheArgThrLeuMetAspValProThr.....ValG1YA 1027
|||
2894 GGTGCAGCAGAGCCATCTCCAGCTGGAACCCGACGAGGAGACTGGGCG 2943
|||
1027 spValArgLysAsnAsnLeuGluIleAspGluIleLysAspGluAsnIle 1043
|||
2944 GCGTCTCAACCCATTCCTCCAGGAGATG.....GACTCTGGACT 2981
|||
1044 LysGluGlyProSerAspSerGluAspIleValPhe..... 1055
|||
2982 GAATATGGCATGGGGAGAGACCAAGAGCTCTTCACCCCTTGTGGA 3031
|||
1056GluGlnThrAspThrAspLeuGlnGluLeuG 1066
|||
3032 CCCAGAGTATACCAAGACGTCCTCCACTGATATAAGATGAATGAGG 3081
|||
1066 lnaIaSer.....MetGlnGlnLeuLeuArgGluGluPro 1077
|||
3082 AATCATCAGCGCGAGCTCTTACTACGCAACTCTTATAGCAAGAAAG 3131
|||
1078 GlyGluGluLysSerGluGluGluGluSerValLeuLysAsnSerAspVa 1094
|||
3132 GCCAAATCATGAGCAAGAAAGATTCGGTGGTA..... 3167
|||
1094 lGluProThrAlaAsnGlyThrAspValAlaAspGluAspAsnPro. 1110
|||
3168AATGTAAACCAACCAACATTCGGCTCTATACGACACACAG 3210
|||
1111 SerSerGluSerAlaLeuAsnGlu 1118

Wed May 15 14:14:01 2002

us-09-783-320-4.rml

Page 8

3211 AATCAGAAAATACAGAAACGAT 3234

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:us-09-221-235-3

seq_documentation_block:

Sequence 3, Application US/09221235

Patent No. 6043040

GENERAL INFORMATION:

APPLICANT: Acton, Susan

TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

FILE REFERENCE: MNT-050

CURRENT FILING DATE: 1998-12-28

EARLIER APPLICATION NUMBER: 09/163,115

NUMBER OF SEQ ID NOS: 15

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 3

LENGTH: 906

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(906)

US-09-221-235-3

alignment_scores:

Quality: 485.50

Ratio: 2.503

Percent Similarity: 74.046

Percent Identity: 38.931

alignement_block:

US-09-783-320-4 x US-09-221-235-3

Align seg 1/1 to: US-09-221-235-3 from: 1 to: 906

```
8 GlnlyslleGlygluGlyserpHecllyslAlalleuVallyserTh 24
  |||||||
112 AAGAAAATGGTGCGGACAAATTGATGAAAGTTATACAGCAGCCTGCT 161
  |||||||
24 rGluAspGlyArgGlnTyValleValleGluIleasnIle...SerArgm 40
  |||||||
162 CTGGATGCAAGACAGCTTCTTAAAGGAGTGCAGATATTGATTAA 211
  |||||||
40 etSerSerlyGluArgGlnGluSerArgArgGluValAlaValleuAla 56
  |||||||
212 TCGATGCCAAAGCAGCTGCTGATTCATCAAGAAATAGATCTTCTTAAG 261
  |||||||
57 AsnMetlysnIleProsnIleValGlnTyArgGlnSerPheGlnGluAs 73
  |||||||
262 CAACCTCAACCAACCAATGTATTAATATATATGATCATTCATTAAGA 311
  |||||||
73 nglySerleuTyTlleValMetAspTyCysGlnGlyGlyAspIleu... 88
  |||||||
312 TAATGAACTTAACATAGTTTGGAACTAGCAAGATCTGCACTTAA 361
  |||||||
89 ....PheLysArgIleAsnAlaGlnTyGlyValleuPheGlnGluAsp 103
  |||||||
362 GAATGATCAAGCATTTTAAAGCAAAAGAG...CTAATTCCTGAAGA 408
  |||||||
104 GlnIleleuAspTPrPheValGlnIleCysleuAlaValleuVallysl 120
  |||||||
409 ACTGTTGGAAAGTATTTGTCAGCTTTCAGTTCAGTTCAGTTCAGTTC 458
  |||||||
120 sAspArglyslleleuHisArgAspIleLysSerGlnAsnIlePheLeuT 137
  |||||||
459 TTCTCGAAGATCATGATAGATATATAAACACCACTTAATGCTTCA 508
  |||||||
137 hrlYsAspGlyThrValGlnleuGlyAspPheGlyIleAlaArgValleu 153
  |||||||
509 CAGCCACTGGGTGTAATAAATTGAGATCTTGGGCTTGCCGCTTTTTC 558
  |||||||
```

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:us-09-221-928-3

seq_documentation_block:

Sequence 3, Application US/09221928

Patent No. 6121030

GENERAL INFORMATION:

APPLICANT: Acton, Susan

TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

FILE REFERENCE: MNT-050

CURRENT FILING DATE: 1998-12-28

EARLIER APPLICATION NUMBER: 09/163,115

NUMBER OF SEQ ID NOS: 15

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 3

LENGTH: 906

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(906)

US-09-221-928-3

alignment_scores:

Quality: 485.50

Ratio: 2.503

Percent Similarity: 74.046

Percent Identity: 38.931

alignement_block:

US-09-783-320-4 x US-09-221-928-3

Align seg 1/1 to: US-09-221-928-3 from: 1 to: 906

```
8 GlnlyslleGlygluGlyserpHecllyslAlalleuVallyserTh 24
  |||||||
112 AAGAAAATGGTGCGGACAAATTGATGAAAGTTATACAGCAGCCTGCT 161
  |||||||
24 rGluAspGlyArgGlnTyValleValleGluIleasnIle...SerArgm 40
  |||||||
162 CTGGATGCAAGACAGCTTCTTAAAGGAGTGCAGATATTGATTAA 211
  |||||||
40 etSerSerlyGluArgGlnGluSerArgArgGluValAlaValleuAla 56
  |||||||
```

```

212 TGGATGCCAAGCAGCTGCTGATTCATCAAGAAATAGATCTTCTTAAG 261
57 AsnMetLysHisProAsnIleValGlnIleValArgLysSerPheGluAsp 73
262 CAACCTCAACACATCAAAATGTAATAAATATATGATCATCATCATTAAGAGA 311
73 nglySerLeuTyrIleValMetAspTyrCysGluGlyValLeuPheGlnAsp 88
312 TAATGAACATAACATAGTTTGGAACTAGACATGCTGCGACCTATCA 361
89 ..... PheLysArgIleAsnAlaGlnIleValLeuPheGlnAsp 103
362 GAATGATCAAGACATTTTAGAAGCAAAAGAG...CTAATTCCTGAAGA 408
104 GlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeuIleValHisValH 120
409 ACTGTTCGAAGATTTTGTTCAGCTTTCACATGATTCATGGAACATCA 458
120 SASPARGLysIleLeuHisArgAspIleLysSerGlnAsnIlePheLeuT 137
459 TTCTCGAAGAGTCATGCATAGAGATATATAAACACAGCTAATGTGTTCA 508
137 hrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAlaArgValLeu 153
509 CAGCCACTGGGCTGTAATAACTTGGAGATCTGGGCTTGGCCGGTTTTC 558
154 AsnSerThrValGlnLeuAlaArgThrCysIleGlyThrProTyrTyrLe 170
559 AGCTCAAAAACACACAGCTGCACATCTTTAGTTGGTACGCCCTTATTAAT 608
170 userProGluIleCysGluAsnLysProTyrAsnAsnLysSerAspIleT 187
609 GCTCCAGAGAGAAATACATGATAAATGGATCAACTTCAATCTGCACATCT 658
187 rPalLeuGlyCysValLeuTyrGlnLeuCysThrLeuLysHisAlaPhe 203
659 GGTCTCTGGCTCTACTATATAGATGCGTCATTAACAAGTCCCTTTC 708
204 GluAlaGlySerMetLys.....AsnLeuValLeuLysIleLeuSer 218
709 TATGCTGACAAAATGAAATTTTACTCTGCTGTAAGAGATAGAACAGTG 758
218 ySerPheProProVal...SerLeuHisTyrSerTyrAspLeuArgSerL 234
759 TGACTACCCACCTCTTCTCTTACATGATCACTATTCAGAAAGATCGCAGCT 808
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
809 TAGTTAATATGTGCATCAACCCAGATCCAGAGAACGACACGACGCTACCC 858
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIle 262
859 TATGTTTATGAC.....GTACCAAAAGAGATG 885

seq_name: /cgn2_6/ptodata/1/lna/6a_comb.seq:us-09-221-527-3
seq_documentation_block:
; Sequence 3, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(906)
US-09-221-527-3

Alignment_scores:
Quality: 485.50 Length: 262
Ratio: 2.503 Gaps: 6
Percent Similarity: 74.046 Percent Identity: 38.931

alignment_block:
US-09-783-320-4 x US-09-221-527-3 ..
Align seg 1/1 to: US-09-221-527-3 from: 1 to: 906

8 GlnIleValLeuGlySerPheGlyLysAlaIleValLysSerThr 24
112 AAGAAATTTGTCGCCGACATTTTACTGAAAGTTTATAGACAGCTGTCT 161
24 rGluAspGlyArgGlnIleValIleLysGluIleAsnIle...SerArg 40
162 CTGGATGAGATACCACTAGCTTATAAAGAGTCAATGATTTGATTTAA 211
40 eSerSerLysGluArgGlnIleValSerArgArgGluValAlaValLeuAla 56
212 TGGATGCCAAGCAGCTGCTGATTCATCAAGAAATATGATCTTCTTAAG 261
57 AsnMetLysHisProAsnIleValGlnIleValArgLysSerPheGluAsp 73
262 CAACCTCAACATCAAAATGTAATAAATATATGATCATCATCATTAAGAGA 311
73 nglySerLeuTyrIleValMetAspTyrCysGluGlyValLeuPheGlnAsp 88
312 TAATGAACATAACATAGTTTGGAACTAGACATGCTGCGACCTATCA 361
89 ..... PheLysArgIleAsnAlaGlnIleValLeuPheGlnAsp 103
362 GAATGATCAAGACATTTTAGAAGCAAAAGAG...CTAATTCCTGAAGA 408
104 GlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeuIleValHisValH 120
409 ACTGTTCGAAGATTTTGTTCAGCTTTCACATGATTCATGGAACATCA 458
120 SASPARGLysIleLeuHisArgAspIleLysSerGlnAsnIlePheLeuT 137
459 TTCTCGAAGAGTCATGCATAGAGATATATAAACACAGCTAATGTGTTCA 508
137 hrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAlaArgValLeu 153
509 CAGCCACTGGGCTGTAATAACTTGGAGATCTGGGCTTGGCCGGTTTTC 558
154 AsnSerThrValGlnLeuAlaArgThrCysIleGlyThrProTyrTyrLe 170
559 AGCTCAAAAACACACAGCTGCACATCTTTAGTTGGTACGCCCTTATTAAT 608
170 userProGluIleCysGluAsnLysProTyrAsnAsnLysSerAspIleT 187
609 GCTCCAGAGAGAAATACATGATAAATGGATCAACTTCAATCTGCACATCT 658
187 rPalLeuGlyCysValLeuTyrGlnLeuCysThrLeuLysHisAlaPhe 203
659 GGTCTCTGGCTCTACTATATAGATGCGTCATTAACAAGTCCCTTTC 708
204 GluAlaGlySerMetLys.....AsnLeuValLeuLysIleLeuSer 218
709 TATGCTGACAAAATGAAATTTTACTCTGCTGTAAGAGATAGAACAGTG 758
218 ySerPheProProVal...SerLeuHisTyrSerTyrAspLeuArgSerL 234
759 TGACTACCCACCTCTTCTCTTACATGATCACTATTCAGAAAGATCGCAGCT 808
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250

```

```

|||||:.....:.....:|||||:.....:
809 TAGTAAATATGTCATCAACCCAGATCCAGAGAAAGCAGCAGCAGTCACC 858
251 SerIleuGluLysGlyPheIleAlaLysArgIle 262
   :|||:.....:
859 TATGTTATATGAC.....GTAGCAAAAGAGATG 885

seq_name: /cgn2_6/plodata/1/lna/6A.COMB.seq:us-09-221-236-3

```

```

seq_documentation_block:
; Sequence 3, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(906)
us-09-221-236-3

```

```

alignment_scores:
Quality: 485.50      Length: 262
Ratio: 2.503        Gaps: 6
Percent Similarity: 74.046      Percent Identity: 38.931

```

```

alignment_block:
US-09-783-320-4 x US-09-221-236-3 ..

```

```

Align seg 1/1 to: US-09-221-236-3 from: 1 to: 906

```

```

8 GlnLysIleGlyGluGlySerPheGlyLysAlaIleLeuValLysSerTh 24
   :|||:.....:
112 AAGAAATATGTCGCGGACAATTAGTGAAGTATATAGACGACGCTGCT 161
24 rGluAspGlyArgGlnTyrValIleLysGluIleAsnIle...SerArg 40
   :|||:.....:
162 CTGGATGAGTACCAAGTACCTTTAAAGGATGACATATTCATTTAA 211
40 eTSerSerLysGluArgGluGluSerArgArgGluValAlaIleVala 56
   :|||:.....:
212 TGATGCCAAAGCAGCTGCTGATTCATCAAGAAATAGATCTTCTTAAG 261
57 AsnMetLysHisProAsnIleValGlnTyrArgGluSerPheGluGlu 73
   :|||:.....:
262 CAACTGACATCAAAATGTAATATATATATGATCATCATTCATGAGAA 311
73 nclYSerLeuTyrIleValMetAspTyrCysGluGlyAspLeu.... 88
   :|||:.....:
312 TATGAACTAAACATAGTTTGGAACTAGCAGATGCTGGCGACCTATCCA 361
89 .....PheLysArgIleAsnIleGlnLysGlyValLeuPheGlnLys 103
   :|||:.....:
362 GAATGATCAACGATTTTAAAGCAAAAGAGG...CTAATTCCTGAAGA 408
104 GlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeuLysHisVal 120
   :|||:.....:
409 ACTGTTGGAAGTATTTGTTGACGCTTGACGTCATTCGAAACATGCA 458
120 sAspArgLysIleLeuHisArgAspIleLysSerGlnAsnIlePheLeu 137
   :|||:.....:
459 TTTCTCAAGATCATCATAGATATATAAACCAGCTAATGTGTCTTCA 508

```

```

137 hTlysAspGlyThrValGlnLeuGlyAspPheGlyIleAlaArgValLeu 153
   :|||:.....:
509 CAGCCATGCGGCTGTAACCTGGAGATCTTGCGCTGGCGGTTC 558
154 AsnSerThrValGlnLeuAlaArgThrCysIleGlyThrProTyrTyle 170
   :|||:.....:
559 AGCTCAAAACCAACAGCTGCATCTCTTAGTTGGACCCCTTATATCAT 608
170 userProGluIleCysGluAsnLysProTyrAsnAsnLysSerAspIle 187
   :|||:.....:
609 GTCGCCAGAGATATCATATAAATGATACAACTTCAATTCGACATCT 658
187 rPalAlaLeuGlyCysValLeuTyrGluLeuGlyThrLeuLysHisAla 203
   :|||:.....:
659 GGCTCTTGCTGCTGCTACTATATGATGATGCTGCTGCTGCTTTC 708
204 GluAlaGlySerMetLys.....AsnLeuValLeuLysIleLeuSerG 218
   :|||:.....:
709 TATGTGACAAATGAATTTATATCTCATCTGTTAGAAAGATGAACAG 758
218 ySerPheProVal...SerLeuHisTyrSerTyrAspLeuArgSerL 234
   :|||:.....:
759 TGACTACCCACCTTCCTTCATCATCATCATTCAGAGAAACCTCCAG 808
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAs 250
   :|||:.....:
809 TAGTTATATGTCATCAACCCAGATCCAGAGAAAGCAGCAGCAGTCAC 858
251 SerIleLeuGluLysGlyPheIleAlaLysArgIle 262
   :|||:.....:
859 TATGTTATATGAC.....GTAGCAAAAGAGATG 885

```

```

seq_name: /cgn2_6/plodata/1/lna/6A.COMB.seq:us-09-221-416-3

```

```

seq_documentation_block:
; Sequence 3, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(906)
us-09-221-416-3

```

```

alignment_scores:
Quality: 485.50      Length: 262
Ratio: 2.503        Gaps: 6
Percent Similarity: 74.046      Percent Identity: 38.931

```

```

alignment_block:
US-09-783-320-4 x US-09-221-416-3 ..

```

```

Align seg 1/1 to: US-09-221-416-3 from: 1 to: 906

```

```

8 GlnLysIleGlyGluGlySerPheGlyLysAlaIleLeuValLysSerTh 24
   :|||:.....:
112 AAGAAATATGTCGCGGACAATTATAGTGAAGTATATAGACGACGCTGCT 161
24 rGluAspGlyArgGlnTyrValIleLysGluIleAsnIle...SerArg 40

```

```

162 CTTGGATGAGTACAGTACAGTCTTTAAAGTCAAGTATTTATTTAA 211
|||||
40 etSerSerlysgluargluserarggluvalalavalleuala 56
|||||
212 TGGATGCCAAGACAGCTGCTGATGCATCAAGAAATGATCTTAAAG 261
57 AsnMetlyshisproasnilevalintyrarggluserphegluuals 73
|||||
262 CACTGACACATCCAAATGTAATTAATATATGATCATTCATTAAGA 311
73 nglyserleutyrllevalmetaspyrcysgluglyaspheleu... 88
|||||
312 TAAAGACTAACAATAGTTTGGACATGACATGCTGGCAGCATCA 361
89 .....PheLysArgIleAsnAlaGlnLysGlyValleuPheGlnLysp 103
|||||
362 GAATGATCAAGCATTTTAAAGACCAAGAG...CTAATTCCTGAAGA 408
104 GlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeuLysHisValH 120
|||||
409 ACTGTTGGAGATATTTTGTTCAGCTTTCAGTGTGGAACACATGCA 458
120 SASPARGLYSIleLeuHisArgAspIleLysSerGlnAsnIlePheLeu 137
|||||
459 TTCTCGAAGAGTCAATGATAGATATTAACCAAGCTAATGCTCATTA 508
137 hLysAspGlyThrValGlnLeuGlyAspPheGlyIleAlaArgValleu 153
|||||
509 CAGCCACGTGGGTGTAACCTTGAGATCTTGCGCTTGCGCTTTTC 558
154 AsnSerThrValGlnLeuAlaArgThrCysIleGlyThrProTyrTyle 170
|||||
559 ACCTCAAAACACAGCTGACATCTTGTAGTGTGACGCTTATATACAT 608
170 userProGlnIleCysGlnAsnLysProTyrAsnAsnLysSerAspIleT 187
|||||
609 GCTCCAGAGAGAAATACATGAATATGATCAACTCAATCTGACATCT 658
187 rPalLeuGlyCysValleuTyrGlnLeuCysThrLeuLysHisAlaPhe 203
|||||
659 GGTCTTGCTGCTGCTACTATATGATGATGCTGATTAACAAAGTCTTC 708
204 GlnAlaGlySerMetLys.....AsnLeuValleuLysIleIleSerG 218
|||||
709 TATGTGACAAATGATATTTATACACGTGTGAAGATAGACAGTG 758
218 ySerPheProVal...SerLeuHisTyrSerTyrAspLeuArgSerL 234
|||||
759 TGACTACCCACCTCTTCCTTCAGATCACTATTCAGAAAGATCCGACAG 808
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
|||||
809 TACTTAATATGTGCATCAACCCAGATCCAGAGAGACGACAGCTCAC 858
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIle 262
|||||
859 TATGTTATATGAC.....GTACCAAGAGAGATG 885

seq_name: /sgn2_6/ptodata/1/lna/6B_COMB.seq:us-09-221-245-3

seq documentation block:
Sequence 3, Application US/09221245
Patent No. 6180358
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMT-050
CURRENT APPLICATION NUMBER: US/09/221,245
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: US 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO: 3
? LENGTH: 906
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(906)
US-09-221-245-3

alignment_scores:
Quality: 485.50 Length: 262
Ratio: 2.503 Gaps: 6
Percent Similarity: 74.046 Percent Identity: 38.931

alignment_block:
US-09-783-320-4 x US-09-221-245-3 ..

Align seg 1/1 to: US-09-221-245-3 from: 1 to: 906
8 GlnLysIleGlyGluGlySerPheGlyValAlaIleLeuValLysSerTh 24
|||||
112 AAGAAATTTGGTCGCGGACATTTAGAGATTTATAGACAGCCTGCT 161
24 rGluAspGlyArgGlnTyrValIleLysGlnIleAsnIle...SerArgm 40
|||||
162 CTGGATGAGTACAGTACAGTCTTTAAAGTCAAGTATTTGATTTAA 211
40 etSerSerlysgluargluserarggluvalalavalleuala 56
|||||
212 TGGATGCCAAGACAGCTGCTGATGCATCAAGAAATGATCTTAAAG 261
57 AsnMetlyshisproasnilevalintyrarggluserphegluuals 73
|||||
262 CACTGACACATCCAAATGTAATTAATATATGATCATTCATTAAGA 311
73 nglyserleutyrllevalmetaspyrcysgluglyaspheleu... 88
|||||
312 TAAAGACTAACAATAGTTTGGACATGACATGCTGGCAGCATCA 361
89 .....PheLysArgIleAsnAlaGlnLysGlyValleuPheGlnLysp 103
|||||
362 GAATGATCAAGCATTTTAAAGACCAAGAG...CTAATTCCTGAAGA 408
104 GlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeuLysHisValH 120
|||||
409 ACTGTTGGAGATATTTTGTTCAGCTTTCAGTGTGGAACACATGCA 458
120 SASPARGLYSIleLeuHisArgAspIleLysSerGlnAsnIlePheLeu 137
|||||
459 TTCTCGAAGAGTCAATGATAGATATTAACCAAGCTAATGCTCATTA 508
137 hLysAspGlyThrValGlnLeuGlyAspPheGlyIleAlaArgValleu 153
|||||
509 CAGCCACGTGGGTGTAACCTTGAGATCTTGCGCTTGCGCTTTTC 558
154 AsnSerThrValGlnLeuAlaArgThrCysIleGlyThrProTyrTyle 170
|||||
559 ACCTCAAAACACAGCTGACATCTTGTAGTGTGACGCTTATATACAT 608
170 userProGlnIleCysGlnAsnLysProTyrAsnAsnLysSerAspIleT 187
|||||
609 GCTCCAGAGAGAAATACATGAATATGATCAACTCAATCTGACATCT 658
187 rPalLeuGlyCysValleuTyrGlnLeuCysThrLeuLysHisAlaPhe 203
|||||
659 GGTCTTGCTGCTGCTACTATATGATGATGCTGATTAACAAAGCTTC 708
204 GlnAlaGlySerMetLys.....AsnLeuValleuLysIleIleSerG 218
|||||
709 TATGTGACAAATGATATTTATACACGTGTGAAGATAGACAGTG 758

```

218 ySerPheProProVal...SerLeuHisTyrSerTyrAspLeuArgSerL 234
 759 TGACACACCACCTCTTCCTTCACATGATGATTAAGAAAGTCCGACAGT 808
 234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
 809 TAGTAAATATGTGCAATCAACCCAGATCCAGAGAGCCAGACAGATCACC 858
 251 SerLeuLeuGluLysGlyPheLeuAlaLysArgTle 262
 859 TATGTTTATGAC.....GTAGCAAAAGAGCATG 885

seq_name: /cgn2_6/ptodata/1/lna/5B_COMB.seq:US-09-163-115-3

seq_documentation_block:

Sequence 3, Application US/09163115A
 Patent No. 6183962
 GENERAL INFORMATION:
 APPLICANT: Acton, Susan
 TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 FILE REFERENCE: NMI-050
 CURRENT APPLICATION NUMBER: US/09/163,115A
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 3
 LENGTH: 906
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(906)
 US-09-163-115-3

alignment_scores:

Quality: 485.50 Length: 262
 Ratio: 2.503 Gaps: 6
 Percent Similarity: 74.046 Percent Identity: 38.931

alignment_block:

US-09-783-320-4 x US-09-163-115-3 ..

Align seg 1/1 to: US-09-163-115-3 from: 1 to: 906

8 GlnTysIleGlyGluGlySerPheGlyLysAlaIleLeuValLysSerTh 24
 112 AAGAAATATGTCGCGGACAAATTAGTGAAGTTATAGACAGCCTGCT 161
 24 rGluAspGlyArgGlnTyrValIleLysGluIleAsnIle...SerArgM 40
 162 CTTGATGAGTACAGTACAGTATTAAAAAAGTCAGATATTTCATTAA 211
 40 eTserSerLysGluArgGluGluSerArgArgGluValAlaValLeuAla 56
 212 TGGATGCCAAAGACGCTGCTGATTCATCAAAAGAAATAGATCTTTAAG 261
 57 AsnMetLysHisProAsnIleValGlnTyrArgGluSerPheGluGluAs 73
 262 CAAGTCACACATCAAAATGATAAATAATATTATTCATCATTCATTAGAGA 311
 73 nGlySerLeuTyrIleValMetAspTyrCysGluGlyLysPleu.... 88
 312 TATGAAACCTAAACATAGTTTGGAACTAGCAGATGCTGGCAGCTATCCA 361
 89PheLysArgIleAsnAlaGlnLysGlyValLeuPheGlnGluAsp 103
 362 GAATGATCAACGATTTTAAAGCAAAAGAGC...CTAATTCCTGAAGA 408
 104 GlnIleLeuAspTrpPheValGlnIleLysLeuAlaLeuLysHisValH 120
 409 ACTGTTTGAAGTATTTGTTTCAGCTTTCAGTTCAGTTCATTCGACACATCA 458

120 sAspArgLysIleLeuHisArgAspIleLysSerGlnAsnIlePheLeuT 137
 459 TTCGCAAGAGTCATGATGATGATTAAGAAAGTCCGACAGT 508
 137 hTlysAspGlyThrValGlnLeuGlyAspPheGlyIleAlaArgValLeu 153
 509 CAGCCACCTGGGCTGTAAACCTTGAGATCTTGCGCTGGCTTTTC 558
 154 AsnSerThrValGluLeuAlaArgThrCysIleGlyThrProTyrTyle 170
 559 AGCTCAAAACACACAGCTGCACATCTTATGTTGAGACCCCTTATATCAT 608
 170 userProGluIleCysGluAsnLysProTyrAsnLysSerAspIleT 187
 609 GTCTCCAGAGAGATATCATATAATGATACAACTTCAATCTGCATCT 658
 187 rPalalaGlyCysValLeuTyrGluLeuGlyThrLeuLysHisAlaPhe 203
 659 GGCTCTTGCTGCTGCTACTATATGATGATGCTGATTCACAAAGTCTTTC 708
 204 GluAlaGlySerMetLys.....AsnLeuValLeuLysIleSerG1 218
 709 TATGTGACAAATGAAATTATATCTACGCTGTAAAGATGAAACAGT 758
 218 ySerPheProProVal...SerLeuHisTyrSerTyrAspLeuArgSerL 234
 759 TGACTACCCACCTTCCTTCACATCATCTATTCAGAAAGTCCGACAGT 808
 234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
 809 TAGTAAATATGTGCAATCAACCCAGATCCAGAGAGCCAGACAGATCACC 858
 251 SerLeuLeuGluLysGlyPheLeuAlaLysArgTle 262
 859 TATGTTTATGAC.....GTAGCAAAAGAGCATG 885

seq_name: /cgn2_6/ptodata/1/lna/6B_COMB.seq:US-09-221-528-3

seq_documentation_block:
 Sequence 3, Application US/09221528
 Patent No. 6190874
 GENERAL INFORMATION:
 APPLICANT: Acton, Susan
 TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 FILE REFERENCE: NMI-050
 CURRENT APPLICATION NUMBER: US/09/221,528
 EARLIER APPLICATION NUMBER: 09/163,115
 EARLIER FILING DATE: 1998-09-29
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 3
 LENGTH: 906
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(906)
 US-09-221-528-3

alignment_scores:
 Quality: 485.50 Length: 262
 Ratio: 2.503 Gaps: 6
 Percent Similarity: 74.046 Percent Identity: 38.931

alignment_block:

US-09-783-320-4 x US-09-221-528-3 ..

Align seg 1/1 to: US-09-221-528-3 from: 1 to: 906

8 GlnTysIleGlyGluGlySerPheGlyLysAlaIleLeuValLysSerTh 24


```

112 AAGAAATTCGTCGCGACATTTAGTAAGTTTATAGACAGCGCTGCT 161
24  rgluasgilyargintyValileugsluleasnile...serargm 40
162 CTGGATGGAGTACACAGTACTTTAAAAAAGTCAGATTTTATGATTAA 211
40  etSerSerlysluargluserarggluValalavalaleuala 56
212 TGGATGCCAAAGCAGCTGCTGATGCATCAAGAAATGATCTTTTAA 261
57  AsnMetlyshisproasnileValgintyargluserpheugluas 73
262 CAACCTCAACCATCCAAATGTAATAATATATATGATCATTCATTAAGA 311
73  nglySerleutyrlilevalmetasptyrcysgluglyaspleu... 88
312 TAATGAACTAAACATAGTTTGGAGTACAGATGCTGCGACCTATCA 361
89  ....PhelysargileasnileaglnlysglyValleupheuglnuas 103
362 GAATGATCAAGCATTTTAAAGACAAAGAG...CTAATTCCTGAAGA 408
104 GlnileuaspttrpPheValglnileCysleuAlaleuYshisValH1 120
409 ACTGTTTGAAGATTTTGTTCAGCTTTCAGTGCATTCGAAACATGCA 458
120 sAsparglyslileuuhisargsprielyserglnasnilepheleut 137
459 TTCTCGAAGAGCATGCAATGATGATATTAACACAGCTAATGCTTCATTA 508
137 hTlyaspglyThValglnleuglyasphelylilealargvalleu 153
509 CAGCCACTGGGTGTAAACTTGAGATCTGGCGCTGGCGGTTTTC 558
154 AsnSerThValglnleuAlaArgThCysIleglyThProtyrTyle 170
559 AACTCAAAAACCCAGACAGTGCATCTTGTAGTGGTACGCTTATACAT 608
170 userProgluilecysgluasnlysprotyrasnasnlyseraspillet 187
609 GTCTCCAGAGCAATATCATGAATGATCAACATTCATAATGCAATCT 658
187 rPalaleuAlCysValleutyrgluleucysThleuYshisAlaphe 203
659 GGCTCTGGCTGCTACTATATAGATGAGTGGCTGATTCACAAAGTCTTTC 708
204 GluAlaGlySerMetlys....AsnleuValleuYshisIleSerGI 218
709 TATGTCGACAAATGATTTATACACTGCTGTAAGAAATAGAACAGTG 758
218 ySerPheProProVal...SerleuH1stySerTyAspPheArgserI 234
759 TGACTACCACTCTTCCTTCAGATCATATTCAAGAACTCGACAGT 808
234 euValSerGlnleupheYsarAsnProArgAspArgProSerValAsn 250
809 TACTTAAATATGTCATCAACCCAGATCCAGAAACGACACAGACGAC 858
251 SerIleuGlnlysglyPheIleAlaYsarGile 262
859 TATGTTATGAC.....GTACCAAGAGGATG 885

seq_name: /cgn2.6/ptodata/1/ina/6B_COMB.seq:US-09-593-553-3
seq_documentation_block:
; Sequence 3, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14

```

```

; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(906)
US-09-593-553-3

alignment_scores:
Quality: 485.50 Length: 262
Ratio: 2.503 Gaps: 6
Percent Similarity: 74.046 Percent Identity: 38.931

alignment_block:
US-09-783-320-4 x US-09-593-553-3 ..
Align seg 1/1 to: US-09-593-553-3 from: 1 to: 906

8  glnylsileglygluglyserpheglylsalaleuVallyserTh 24
112 AAGAAATTCGTCGCGACATTTAGTAAGTTTATAGACAGCGCTGCT 161
24  rgluasgilyargintyValileugsluleasnile...serargm 40
162 CTGGATGGAGTACACAGTACTTTAAAAAAGTCAGATTTTATGATTAA 211
40  etSerSerlysluargluserarggluValalavalaleuala 56
212 TGGATGCCAAAGCAGCTGCTGATGCATCAAGAAATGATCTTTTAA 261
57  AsnMetlyshisproasnileValgintyargluserpheugluas 73
262 CAACCTCAACCATCCAAATGTAATAATATATATGATCATTCATTAAGA 311
73  nglySerleutyrlilevalmetasptyrcysgluglyaspleu... 88
312 TAATGAACTAAACATAGTTTGGAGTACAGATGCTGCGACCTATCA 361
89  ....PhelysargileasnileaglnlysglyValleupheuglnuas 103
362 GAATGATCAAGCATTTTAAAGACAAAGAG...CTAATTCCTGAAGA 408
104 GlnileuaspttrpPheValglnileCysleuAlaleuYshisValH1 120
409 ACTGTTTGAAGATTTTGTTCAGCTTTCAGTGCATTCGAAACATGCA 458
120 sAsparglyslileuuhisargsprielyserglnasnilepheleut 137
459 TTCTCGAAGAGCATGCAATGATGATATTAACACAGCTAATGCTTCATTA 508
137 hTlyaspglyThValglnleuglyasphelylilealargvalleu 153
509 CAGCCACTGGGTGTAAACTTGAGATCTGGCGCTGGCGGTTTTC 558
154 AsnSerThValglnleuAlaArgThCysIleglyThProtyrTyle 170
559 AACTCAAAAACCCAGACAGTGCATCTTGTAGTGGTACGCTTATACAT 608
170 userProgluilecysgluasnlysprotyrasnasnlyseraspillet 187
609 GTCTCCAGAGCAATATCATGAATGATCAACATTCATAATGCAATCT 658
187 rPalaleuAlCysValleutyrgluleucysThleuYshisAlaphe 203
659 GGCTCTGGCTGCTACTATATAGATGAGTGGCTGATTCACAAAGTCTTTC 708
204 GluAlaGlySerMetlys....AsnleuValleuYshisIleSerGI 218

```

```

709 TATGTTGACAAATGATTTACTCTGCTAGCAAGATGACAGT 758
      ::::::::::::::::::::
218 ySerPheProProl...SerLeuHisTyrSerTyrAspLeuArgSerI 234
      ::::::::::::::::::::
759 TGACTACCCACCTCTCTCTCTAGATCTATTCAGAAAGCTCCGACGT 808
      ::::::::::::::::::::
224 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
      ::::::::::::::::::::
809 TAGTTAAATATGTCATCAACCCAGATCCAGAGAACGACCAAGCTCAGC 858
      ::::::::::::::::::::
251 SerIleuGlnLysGlyPheIleAlaLysArgIle 262
      ::::::::::::::::::::
859 TATGTTTATGAC.....GTAGCAAGAGGATG 885
      ::::::::::::::::::::

```

seq_name: /cgn2.6/ptodata/1/lna/6B.COMB.seq:US-09-221-237-3

```

seq_documentation_block:
: Sequence 3, Application US/09221237
: Patent No. 6214597
: GENERAL INFORMATION:
: APPLICANT: Accion, Susan
: TITLE OF INVENTION: NOVEL CSAKR-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
: FILE REFERENCE: NMI-050
: CURRENT APPLICATION NUMBER: US/09/221,237
: EARLIER FILING DATE: 1998-12-28
: EARLIER APPLICATION NUMBER: 09/163,115
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 906
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(906)
: US-09-221-237-3

```

alignment_scores:

Quality:	485.50	Length:	262
Ratio:	2.503	Gaps:	6
Percent Similarity:	74.046	Percent Identity:	38.931

alignment_block:

US-09-783-320-4 x US-09-221-237-3

Align seg 1/1 to: US-09-221-237-3 from: 1 to: 906

```

8 GlnLysIleGlyGlySerPheGlyLysAlaIleLeuValLysSerTh 24
      ::::::::::::::::::::
112 AAGAAATTTGGTCGGGACATTTAGTGAAGTTATAGACAGCCTGTCT 161
      ::::::::::::::::::::
24 GlnAspGlyArgGlnTyrValIleLysGlnIleAsnIle...SerArg 40
      ::::::::::::::::::::
162 CTGCGATGAGTACCAGTACCTTTAAAAAAAGTCAGATATTGATTGAA 211
      ::::::::::::::::::::
40 eLysSerLysGlnArgGlnLysSerArgArgGlnValAlaLysAla 56
      ::::::::::::::::::::
212 TGGATGCCAAAGCAGCTGCTGATTCATCAAGAAATAGATCTTCTTAA 261
      ::::::::::::::::::::
57 AsnMetLysHisProAsnIleValIleGlnTyrArgGlnSerPheGln 73
      ::::::::::::::::::::
263 CAACCTCAACCATCCAAATGTAATAATATTATGATCATTCATTGAAGA 311
      ::::::::::::::::::::
73 nGlySerLeuTyrIleValMetAspTyrGlyGlnGlyLysAspLeu... 88
      ::::::::::::::::::::
312 TAAATGACTAAACATAGTTTGGAACTAGACGATGCTGGCGACCTATCCA 361
      ::::::::::::::::::::
89 ....PheLysArgIleAsnAlaGlnLysGlyValLeuPheGlnGlnAsp 103
      ::::::::::::::::::::
363 GAATGATCAAGCATTTTAAAGCAAAAGAGG...CTAATTCTCTGAAGA 408
      ::::::::::::::::::::

```

```

104 GlnIleLeuAspTrpPheValGlnIleCysLeuAlaLysHisValHis 120
      ::::::::::::::::::::
409 ACCTTTGGAAGATATTGTTGTCAGCTTTGACAGTTCGGAACACATGCA 458
      ::::::::::::::::::::
120 sAspArgLysIleLeuHisArgAspIleLysSerGlnAsnIlePheLeu 137
      ::::::::::::::::::::
459 TTCTCAAGAGTATGATCATGAGATATAAACACCACTAATGCTTCATTA 508
      ::::::::::::::::::::
137 hTrpAspGlyThrValGlnLeuGlyAspPheGlyIleAlaArgValLeu 153
      ::::::::::::::::::::
509 CAGCCACTGGGGTGTAACCTTGAGATCTTGCGCTGGCCGGTTTTC 558
      ::::::::::::::::::::
154 AsnSerThrValGlnLeuAlaArgThrCysIleGlyThrProTyrTrpLe 170
      ::::::::::::::::::::
559 AGCTCAAAAACACACAGCTGCACATTCCTTTAGTGTGACCCCTTATACAT 608
      ::::::::::::::::::::
170 uSerProGlnIleCysGlnAsnLysProTyrAsnAsnLysSerAspIle 187
      ::::::::::::::::::::
609 GTCTCCAGAGAGATATACATGAATAATGATACAACTTCAATCTGACATCT 658
      ::::::::::::::::::::
187 rPalLeuGlyCysValLeuTyrGlnLeuCysThrLeuLysHisAlaPhe 203
      ::::::::::::::::::::
659 GGCTCTTGCTGCTCTATATATGAGATGCTGCATTCACAAAGCTCTTTC 708
      ::::::::::::::::::::
204 GlnAlaGlySerMetLys.....AsnLeuValLeuLysIleIleSerGln 218
      ::::::::::::::::::::
709 TATGTCACAAATGATTTATCTACTCTGCTGTAAGAAAGATGAAACAGTG 758
      ::::::::::::::::::::
218 ySerPheProProl...SerLeuHisTyrSerTyrAspLeuArgSerI 234
      ::::::::::::::::::::
759 TGACTACCCACCTCTCTCTCTAGATCTATTCAGAAAGCTCCGACAGT 808
      ::::::::::::::::::::
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
      ::::::::::::::::::::
809 TAGTTAAATATGTCATCAACCCAGATCCAGAGAACGACCAAGCTCAGC 858
      ::::::::::::::::::::
251 SerIleuGlnLysGlyPheIleAlaLysArgIle 262
      ::::::::::::::::::::
859 TATGTTTATGAC.....GTAGCAAGAGGATG 885
      ::::::::::::::::::::

```


Wed, May 15 14:14:01 2002

us-09-783-320-4.rn1

THIS PAGE BLANK (USPTO)

1 ATGGAGAGTATGTTAGACTACAGAAAGATTGGAGAAAGCTTCATTGGAAA 50
17 sAlaIleuValIySerThrGluAspGlyArgInIyValIleIySG 34
|||||
51 AGCCATCTCTGTTAAATCTACAGAAAGATGCGACAGTATGTTATCAAG 100
34 IuleAsnIleSerArgMetSerSerIySGIuArgGluGluSerArg 50
|||||
101 AAATTAAACATCTCAAGATCTCCAGTAAAGAAAGAAACATCAAGAGA 150
51 GluValAlaValIleuAlaAsnMetIySHISProAsnIleValGlnIyR 67
|||||
151 GAAGTTCAGATTGGCAACATGAAAGCATCCAAATATTGTCAGATAG 200
67 gGIuSerPheGluGluAsnGlySerLeuTyrIleValMetAspIyCYSG 84
|||||
201 AGAATCTTTGAAGAAATGCTCTCTACATAGTAATGATGACTGTG 250
84 IuGIyGIyAsPheIyPheIyArgIleAsnAlaGlnIySGIyValIeuPhe 100
|||||
251 AGGAGGCGATCTGTTAAGCGAATAATGCTCAGAAAGCGCTTTGTT 300
101 GlnGluAspGlnIleLeuAspTrpPheValGlnIleCYsIleuAlaIeuIy 117
|||||
301 CAAGAGCATCAGATTTGGACTGTTTGTACAGATATGTTGCCCTGAA 350
117 sHISValIleAspArgIyIleIleuHISArgAspIleIySerGlnAsnI 134
|||||
351 ACATGTCATGATAGAAATAATCTTCATCGACATTAATCTCAGAAC 400
134 IepheLeuThrIyAspGIyThrValGlnIeuGIyAspPheGIyIleAla 150
|||||
401 TATTTTAACTAAAGATGACAGACATCAACTTGCAGATTTTGCATTC 450
151 ArgValIleuAsnSerThrValGluIleuAlaArgThrCYsIleGIyThrP 167
|||||
451 AGAGTTCTTAATAGTACTGTAGAGCTGCTCGAACTTGATGAGGAGCCC 500
167 oTYrTYrLeuSerProGluIleCYsGluAsnIyProTYrAsnAsnIyS 184
|||||
501 ATACTACTTGTCACTGAATCTGTGTAACAAACCTTACAAATAATTA 550
184 eRAspIleTrpAlaIeuGIyCYsValIleuIyGluIeuCYsThrLeuIyS 200
|||||
551 GTGACATTTGGCTGGGCTGTGCTTATGAGCTGTGATCACTTAA 600
201 HIsAlaPheGluAlaGlySerMetIyAsnLeuValIleuIySIIeIese 217
|||||
601 CATCTTTTGAAGCTGGCAGTATGAAACCTGTACTGACATTAATATC 650
217 rGIySerPheProValSerLeuHISArgIyAspLeuArgSerI 234
|||||
651 TGGATCTTTTCCACCTGTCTTGTGATATTCCTATGATCCGCCAGTT 700
234 euValSerGlnIeuPheIyArgAsnProArgAspArgProSerValAsn 250
|||||
701 TGGTGCTAGTATTATAAGAAATCTAGGATAGACATCAGCAAC 750
251 SerIleuGIySGIyPheIleAlaIyArgIleGIyIySGIyPheIeuSe 267
|||||
751 TCCATATTGGAGAAAGTTTATAGCCAAACGATTGAAAGCTTCTCTC 800
267 rPheGlnIeuIleAlaGluIuPheCYsLeuIySThrPheSerIyPheG 284
|||||
801 TCCACACCTATTGACAGAAATTTGTCTAAATAACATTTTGAAGTTG 850
284 IySerGlnProIleProAlaIyArgProAlaSerGIyGlnAsnSerIle 300
|||||
851 GATCAGACCTATACACGTAATAAGACCACTTCAAGCAAAACCTGAT 900
301 SerIleuMetProAlaGlnIySIIeThrIySProAlaIleAlaIySIIy 317
|||||
901 TCTGTATGCTGCTCAGAAATTAACAAACCTGCGCTAAATATAGAAAT 950

317 eProIleuAlaIyRlyIySIIyTYrGIyAspIySlyLeuHISGluIyIyS 334
|||||
951 ACCTTAGCATATTAAAGAAATATGAGATTAATAATTCACGAAAGAAAC 1000
334 rIeuGlnIySHISlySGIuAlaHISGlnIyThrProGluIySArgValAsn 350
|||||
1001 CACTGCAAAACATTAACAGCGCCCATCAACTCCAGAGAAAGAGTGAAT 1050
351 ThrGIyGluGluArgArgIySIIeSerGluIuAlaIleArgIySArg 367
|||||
1051 ACTGGAGAAAGAAAGAGAAATATCTGACGACACAGCAAGAAAGAAAG 1100
367 gIeuGluPheIleGIyIySGIyIySGIyGlnIySAspGlnIleIeSerI 384
|||||
1101 GCTGAATTTATTGAAAGAAAGAAAGAAAGAAAGATCAGATTATAGTT 1150
384 euMetIySAlaGluIuMetIySArgGlnGluIySGIuArgIeuGluArg 400
|||||
1151 TAATGAAGCTGACAAATGAAAGAAAGCAAGAAAGAAAGGTTGAAAGA 1200
401 IleAsnArgAlaArgGluGlnGlnIyTrpArgAsnValIeuSerAlaGIy 417
|||||
1201 ATTAATAGGCGCCAGGAAACAAAGATGAGAAATGCTAAGTCTGTG 1250
417 ySerGIyGluValIySAlaPheIeuGIySerGIyGIyThrIleAlaP 434
|||||
1251 AAGTGTGAAGTAAAGGCTCTTTCTGTGCGCAGTGAAGGACATTAAGCTC 1300
434 rIeSerPheSerSerArgGIyGlnIyRGIyHISArgIleAlaIlePhe 450
|||||
1301 CATCACTTTTCTTCTCTGAGGACAGTATGAACATTAACATGCCATTTT 1350
451 AspGlnMetGlnGluIuArgAlaGluIuAspGlnIuAlaIySArgIyS 467
|||||
1351 GACCAATGACAGCAAGAGACAGAAATATGAGCTTAATGAAAG 1400
467 gGIuIleTYrGIyArgIyIeuProGluArgGlnIySGIyGlnIeuAla 484
|||||
1401 AGAAATATATGTGTCGAGCTTCACGAAAGCAAAAGGAGCTACTG 1450
484 aIGluArgAlaIySGIuValGluIuPheIeuGlnIyArgIuAla 500
|||||
1451 TAGAAAGAGCTTAACAGTAAAGAGATTCCTGACGCGAAACGGGAAGCT 1500
501 MetGlnAsnIySAlaArgAlaGluIuGlnIyHISMetValTYrIleuAlaArg 517
|||||
1501 ATGCAGAAATTAAGCTCGACCGAAGACATATGCTTATCTGCAAGACT 1550
517 uArgGlnIleArgIeuGlnAsnPheAsnGluArgGlnIleIySAlaI 534
|||||
1551 GAGCGAAATTAAGACTAGAAATTTCAATGAGCGCAACAGATTAAAGCA 1600
534 yIleuArgGIyGluIySGIyIuAlaAsnHISSerGIyGluIySGIyG 550
|||||
1601 AACTTCTGTGTAAGAAAGAAAGAAAGCTAATCATTCGAAGGCAAGAAAGA 1650
551 SerGIuIuAlaAspMetArgArgIySIIeGIySerLeuIySAlaHIS 567
|||||
1651 AGGAGAGAGGCTGACATGAGGCGCAAAAAAATCGAATCTAGGAAGCCCA 1700
567 sAlaAsnAlaArgAlaIleValIleuIySGIuIuArgIySArgI 584
|||||
1701 TGCAAATGCACTGCTGCTGTACTTAAAGAAACATCAACAGAAAGAA 1750
584 ySGIuAlaTYrGIuArgGIyIySlySValIyTPGluGlnHISIleuValAla 600
|||||
1751 AGGAGGCTTATGAGAGAAAAAAGTGTGGAAAGACATTTGTGCTGCT 1800
601 IySGIyValIySerSerAspValSerProIeuGlnIyGlnHISGIyuth 617
|||||
1801 AAAGAGTTTAAGATTCTGATGTTTCTCCACCTTTGGGACAGCATGAAC 1850

```

617  rGlyGlySerProSerLysGlnGlnMetArgSerValIleSerValThrS 634
    |||
1851  AGGTGGCTCTCATCAAGCAAGATGATCTGTATTCTCTGAACCTT 1900
634  eValLeuLysGlnValGlyValAspSerSerLeuThrAspThrArgGlu 650
    |||
1901  CAGCTTTGAAAGAAAGTTGGCGTGGACAGTAGTTTAAGTATGATCCGGGAA 1950
651  ThrSerGlnGlnMetGlnLysThrAsnAsnAlaIleSerSerLysArgG1 667
    |||
1951  ACTTCAGAGAGATGCCAAAAGCAACAAATGCTATTCAAGAAAGTGAAGGAGA 2000
667  uIleLeuArgArgLeuAsnGlnLysAsnLeuLysAlaGlnGlnAspGluLysG 684
    |||
2001  AATAGCTGACAGATTAAATGAAATCTTAAAGCTCAAGAAAGATGAAGAAAAG 2050
684  lMetGlnAsnLeuSerAspThrPheGluIleAsnValIleGlnLysAspAla 700
    |||
2051  GAATGACAGAAATCTCTGATACCTTGTGAGTAAATGTTCAATGACAGATGCC 2100
701  LysGlnHisGlnLysGlnLysSerValSerSerAspArgLysLysTrpG1 717
    |||
2101  AAAGAGCATGAAAAAGAAAAATCAGTTTCATCTGATCGCAAGAGTGGGA 2150
717  uAlaGlyGlnGlnLeuValIleProLeuAspGluLeuThrLeuAspThrS 734
    |||
2151  GGCAGAGAGTCACCTTGATTCCTCTGAGTGAATGATTAACACTGATACAT 2200
734  ePheSerThrThrGlnArgHisThrValGlyGlnValIleLysLeuGly 750
    |||
2201  CCTCTCTACACACTGAAGACATACACTGGGAGAAATTAATTAATAGCT 2250
751  ProAsnGlySerProArgArgAlaTrpGlyLysSerProThrAspSerVa 767
    |||
2251  CCTAATGAGATCTCCAAGAGAGCCGGGGGAAAAAGTCCGACAGATCTCTG 2300
767  lLeuLysIleLeuGlnGlnValGlnLeuGlnLeuGlnIleThrGluLeuGln 784
    |||
2301  TCMAAGATCTTGGAGAAAGCTGAACCTCAACTCAGACAGAACTATTAG 2350
784  lAsnThrThrIleArgSerGluIleSerProGlnGlnLysLysLysLys 800
    |||
2351  AAAATACAACTATTAGAGTGAATTTCTCCCGAAAGGGAAGAAAGTACAAA 2400
801  ProLeuIleThrGlyLysLysValGlnCysIleSerHisGluIleAs 817
    |||
2401  CCCTTAATTACTGGAGAAAAAAAGTACAAATGATTTCTCAGTGAATATAA 2450
817  nProSerAlaIleValAspSerProValGluThrLysSerProGluPheS 834
    |||
2451  CCCATCAGCTATTGTTGATTCTCCCTGTGAGACAAAAGTCCGAGTTCA 2500
834  eArgLysIleSerProGlnMetSerLeuLysLeuGlnGlyAsnLeuGlnGlu 850
    |||
2501  GTGAGGCAATCTCCACAGATGTCATGAAACTGGAGAAATTTAGAGAA 2550
2551  ProAspArgLeuGlnThrGlnIleLeuGlnGlnProSerGlyThrAsnLys 867
    |||
867  sAspGlnSerLeuProCysThrIleThrAspValTrpIleSerGlnGluL 884
    |||
2601  AGATGAGAGCTTCCATGCATATTAATCTGATGTGTGATTTAGTGGGAAA 2650
884  LysGlnThrLysGlnThrGlnSerAlaAspArgIleThrIleGlnLysn 900
    |||
2651  AAGAAACAAAGAAAGAACTCAGTCCGCGCATAGGATCAACCATTCAGAAAT 2700
901  GluValSerGlnAspGlyValSerSerThrValAspGlnLeuSerAspI1 917
    |||
2701  GAAGTTCTGAGAGTGAAGTCTGAGTACTGTGACCAACTTAGGACAT 2750
917  eHisIleGlnProGlyThrAsnAspSerGlnHisSerLysCysAspValA 934

```

```

2751  TCATATAGAGCTGGAGAACCATATTCTCAGCACTCTAATATGATGATAG 2800
934  sPlySerValGlnProGlnProPhePheHisLysValValHisSerGlu 950
    |||
2801  ATAACTGTGCAACCGGAACCATTTTTCATTAAGTGGTTCATTCGAA 2850
951  HisLeuAsnLeuValProGlnValGlnSerValGlnCysSerProGlnG1 967
    |||
2851  CACTTGAACCTTAGTCCCTCAAGTTCATTCAGTTCATTCACGAAAGA 2900
967  uSerPheAlaPheArgSerHisSerHisLysLeuProPlyAsnLysAsnL 984
    |||
2901  ATCCCTTCATTTCCATCTCAGTCCATGCAATTAACCAAAAATAAACA 2950
984  yAsnSerLeuLeuIleGlyLeuSerThrGlyLeuPheAspAlaAsnAsn 1000
    |||
2951  AGAATTCCTGCTGATTCGACTTTCACCTGCTGTTGATGCAAAACAC 3000
1001  ProLysMetLeuArgThrCysSerLeuProAspLeuSerLysLeuPheAr 1017
    |||
3001  CCAAGATGTTAAGACATGTTTCACTTCAGATCTCTCAAAAGCTGTTCC 3050
1017  gThrLeuMetAspValProThrValGlyAspValArgGlnAspAsnLeuG 1034
    |||
3051  AACCTTATGATGATGTTCCACCGTAGAGATGTTCCGACAGAACATCTG 3100
1034  lValLeuAspGluIleLysAspGlnAsnIleLysGlnGlnProSerAspSer 3150
    |||
3101  AAATAGATTAATTAAGATGAAACATTAAGAAAGACCTCTGATTCCT 3150
1051  GluAspIleValPheGlnGlnIleThrAspThrAspLeuGlnLeuGlnAl 1067
    |||
3151  GAAGACATTTGTTGTGAAACAGTACACAGATTTTACAAAGCTGCAGGC 3200
1067  aSerMetGlnGlnLeuLeuArgGlnGlnProGlyGlnGlnLysSerGln 1084
    |||
3201  CTCGATGGAACAGTTACTTAGGGAACACCTGGTGAAATACAGTGAAC 3250
1084  lGlnGlnLeuSerValLeuLysAsnSerAspValGluProThrAlaAsnGly 1100
    |||
3251  AAGAAAGTGCAGCTTGAAGAACAGTGAATGTGGACCAACTGCAAAAGGG 3300
1101  ThrAspValAlaAspGlnAspAsnProSerSerGlnSerAlaLeuAs 1117
    |||
3301  ACAGATCTGCAGATGAAAGATGACAAATCCAGTATGTGAAGTCCCTGAA 3350
1117  nGlnGlnTrpHisSerAspAsnSerAspGlyGlnIleAlaSerGluCysG 1134
    |||
3351  CGAAAGATGGCACTCAGATTAACAGTGAATGTGAATTTCTAGTGAATGTG 3400
1134  lCysAspSerValPheAsnHisLeuGlnGlnLeuArgLeuHisLeuGln 1150
    |||
3401  AATCGATAGTGTCTTTAACCATTTAGAGAACCTGAGACTCAGCTGGAG 3450
1151  GlnGlnMetGlyPheGlnLysPhePheGlnValIlyGlnLysIleLysAl 1167
    |||
3451  CAGAAATGGCTTTGAAAAATTTCTTGAGTTTGAAGAAATTAAGAGGC 3500
1167  aIleHisGlnAspGlnAspGlnAsnIleGlnLysSerLysIleValG 1184
    |||
3501  TATTTCATGAAGATGAAATGAAATATTAATAATTTGTTCCAAAAATAGTTC 3550
1184  lAsnIleLeuGlnGlnAsnGlnHisGlnHisLeuTrpAlaLysIleLeuHis 1200
    |||
3551  AAAATATTTTGGGAAAGAACATCAGCATCTTAATGAGCAATCTTTCAT 3600
1201  LeuValMetAlaAspGlyAlaTrpGlnGlnAspAsnAspGlu 1214
    |||
3601  TTAGTCATGGCGCATGTGAGGCTTCCAAAGAAAGTAAATGATGAAG 3642

```

seq_name: /net/abs06/s1db1/9cgcdata/hold_geneseq/geneseq-emb1/NA2001A.DAT:AA511582

seq_documentation_block:
 ID AAS11582 standard; cDNA: 5426 BP.
 AC AAS11582;
 DT 24-OCT-2001 (first entry)
 DE Human cDNA encoding novel human protein, NHP #26.
 XX
 KW Human: novel human protein; NHP; ss; breast cancer; prostate cancer;
 immunogen; antibody; gene therapy; antisense.
 XX
 OS Homo sapiens.
 PN W02001.61016-A2.
 XX
 PD 23-AUG-2001.
 XX
 PE 15-FEB-2001; 2001MO-US05356.
 XX
 PR 18-FEB-2000; 2000US-0183582.
 PR 22-FEB-2000; 2000US-0184014.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Hu Y, Nepomichy B, Turner CA, Zambrowicz B;
 DR WPI; 2001-502793/55.
 XX
 PT Isolated nucleic acids encoding novel human proteins useful for the
 PT treatment of disease and as probes for testing and detection -
 XX
 PS Disclosure: Page 68-69; 69pp; English.
 XX
 CC The invention relates to novel human proteins (NHP) and the nucleic
 CC acids encoding them. The nucleic acids encode mammalian transporter
 CC proteins and are useful for the treatment (e.g. by gene therapy or
 CC antisense technology) of any of a wide variety of symptoms associated
 CC with biological disorders (e.g. breast and prostate cancer) or imbalances
 CC and as probes for the identification, selection and validation of novel
 CC molecular targets for drug discovery. The proteins may be used to raise
 CC anti-NHP antibodies. The present sequence encodes an NHP of the
 CC invention.
 XX
 SO Sequence 5426 BP; 1852 A; 927 C; 1163 G; 1484 T; 0 other;

alignment_scores:
 Quality: 6243.00 Length: 1214
 Ratio: 5.143 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

Alignment block:
 US-09-783-320-4 x AAS11582 ..

Align seg 1/1 to: AAS11582 from: 1 to: 5426

```

1 MetGluGlySerValArgLeuGlnLysIleGlyGluGlySerPheGly 17
  |||||
534 ATGGCAAGATGTTAGACTACAGAACTTGGAGAGAGTTCATTGGAAA 563
  |||||
17 salalleuvallyserthrGluaspGlyArgGlnTyrValIleLysG 34
  |||||
584 ACCCATCTCTGTTAAATCTACAGAAAGATGCGACAGTATGTATTCAGG 633
  |||||
34 luileasnlleSeraTgmetSerSerLysGluATgGluGluSerArgArg 50
  |||||
634 AAATTACATCTCCAAAGATGTCAGTAAAGAAAGAAAGATCAAGGAGA 683
  |||||
51 GluValAlaValIleuAlaAsnMetLysHisProAsnIleValGlnTyr 67
  |||||
684 GAAGTTGCGATGTGGCAAAACATGAAGCATCAATATATTGTCAGTATAG 733
  |||||

```

```

67 gGluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
  |||||
734 AGAATTCATTTGAGAAATATGGCTCTCTACATAGTAATGATGTTCTGTG 783
  |||||
84 luGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
  |||||
784 AGCGAGCGGATCTGTTTAAACCAATTAATGCTCAGAAAGCCCTTTGTTT 833
  |||||
101 GlnGluAspGlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeuLys 117
  |||||
834 CAAGAGATCAGATTTGGACTGTTGTACAGATATGTGTTGGCCCTGAA 883
  |||||
117 SHISValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsnI 134
  |||||
884 ACATGTACATGATAGAAAAATTTCTCATGACACATTAATCTCAGAAC 933
  |||||
134 lePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
  |||||
934 TATTTTAACTTAAAGATGAAACAGTACAACTTGAGATTTTGGAATGCT 983
  |||||
151 ArgValLeuAsnSerThrValGluLeuAlaArgThrCysIleGlyThrPr 167
  |||||
984 AGAGTTCTTAATAGTACTGTAGAGCTGGCTCGAACTTGATGGGAGCC 1033
  |||||
167 OTYrTYrLeuSerProGlnIleCysGluAsnLysProTyrAsnAsnLys 184
  |||||
1034 ATACTACTGTGCACCTGAATCTGTGAAAAACAACCTTACAAATTAATAA 1083
  |||||
184 eAspIleTrpAlaLeuGlyCysValLeuTyrGluLeuCysThrLeuLys 200
  |||||
1084 GTGACATTTGGGCTCTGGGGTGTGCTCTTATAGAGCTGTGACTACTTAA 1133
  |||||
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleIleS 217
  |||||
1134 CATCTTTTGAAGCTGGCACTATGAAAAACCTGTACTTAACATATATATC 1183
  |||||
217 rGlySerPheProProValSerLeuHisTyrSerThrAspLeuArgSerL 234
  |||||
1184 TGGATCTTTCACCTGTGTGCTTTCATATTCCTATGATCTCCCAAGTT 1233
  |||||
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
  |||||
1234 TGGTGTCTCAGTATTTAAAGAAATCTTAGGATAGACATCACTCACTAC 1283
  |||||
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGluLysPheLeuS 267
  |||||
1284 TCCATATTTGGAGAAAGCTTTTATAGCCAAACGATTTGAAAGTTTCTCTC 1333
  |||||
267 rProGlnLeuIleAlaGluGluPheCysLeuLysThrPheSerLysPheG 284
  |||||
1334 TCCCTACGCTTATTCAGAAAGAAATTTGTCTAAAAACATTTTCGAGATTG 1383
  |||||
284 LysSerGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
  |||||
1384 GATCACAGCCTATATCCAGTAAAGAACAGCTTCAGAGCAAAAAACGATTT 1433
  |||||
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGlyT 317
  |||||
1434 TCTGTTATGCTGCTCAGAAATTTACAAAGCCTGCCGCTTAATATAGCAT 1483
  |||||
317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGluLysLysP 334
  |||||
1484 ACCTTTAGCATATTAAGAAATATGAGATTAATAAAATTAACCGAAAAAGAAC 1533
  |||||
334 rLeuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
  |||||
1534 CACTGCAAAAAACATAAACAGGCCCATCAAACTCCAGAGAGAGAGATGAT 1583
  |||||
351 ThrGlyGluGluArgArgLysIleSerGluGluAlaAlaArgLysArgAr 367
  |||||
1584 ACTGAGAGAGAAAGAGAAATATATCTGAGGAAGCGCAAGAAAGAGAGAG 1633
  |||||
367 gLeuGluPheIleGlnLysGlnLysLysGlnLysAspGlnIleIleSerL 384
  |||||

```

```

1634 GCTGCAATTTATGAAAAAGAAAAAGAAAGATCATATTATAGTT 1683
1634 eumellysAlagInglumetylsArgIngluysgluArleugluAr 400
1684 TAATGAAGGCTGACACAAATGAAAAAGCAAGAAAGGTTGGAAAG 1733
401 tleasnArAlarglugIngluYrParqasnValleuserlAglYgl 417
1734 ATAAATAGGCGCAAGCAAGAGATGGAAATAGCTAGTCTGCTGG 1783
417 ySerGluVallylsAlaProPheLeuGlySerGlyThrIleAla 434
1784 AACTGTGAAGTAAAGGCTCTTTCTGTGGCAAGGAGACATATACCT 1833
434 roserSerPheSerSerArglyIngluYrGluInlYrHisAlaIlePhe 450
1834 CATCATCTTTCTCTCGAGCAGATATGAACATTCACATGCAATTTT 1883
451 AspGlnmetGIngluInArGAlaGluaspasnGluAlaYstrPlySar 467
1884 GACCAATGCGCACCAAGACAGACAGATATGAAGCTAAATGCAAAAG 1933
467 gGluIleTyrGlyArgGlyLeuProGluArGluInlYsGlyInleuAla 484
1934 AGAAATATATGTCGAGGCTTCACAAAGGCAAAAGGCAAGCTAGCTG 1983
484 alGluArGAlaYsGluInValGluPheLeuGluInArGlyArGluAla 500
1984 TAGAAAGCTAAACAAAGTAGAAGATTCTCTGACGAAAGGGAACCT 2033
501 MetGlnsnlylsAlaArGAlaGluGluYrHisMetValTyrLeuAlaAr 517
2034 ATGCAGATTAAGCTCGAGCCGAAAGCATATGTTATCTGCAAGACT 2083
517 uArGlnIleArGleuGlnsnPheasnGluArGIngluInleYsAla 534
2084 GAGCAATTAAGCTACAGAAATTCATGAGCCCAACAGATTAAAGCA 2133
534 ySleuArGlyGluLysGluAlaasnHisSerGluIngluYgl 550
2134 AACCTCGTGTGAAAAGAAAGAAAGTATCTGAGGACCAAGAGCA 2183
551 SerGluGluAlaaspMetArGArGlyLysIleGluSerleuYsAlaH 567
2184 AOTGAAGAGCTGACATGAGGCGCAAAATGCAATCTGAGAGCCCA 2233
567 sAlaasnAlaArGAlaAlaValleuYsGluInleuGluArGlyAr 584
2234 TCCAATGCGAGCTGCTCTCTACTTAAAGAACTAGAACGAAAGAA 2283
584 ySgluAlaTyrGluArGluLysLysValTyrGluInlHisleuAla 600
2284 AGGAGCTTATGAGAGAGAAAAAAGTGTGGAAAGCAATTTGGTGGCT 2333
601 lysGlyVallylsSerSerAspValSerProPheGluInlHisGlu 617
2334 AAGGAGTTAAGAGTTCTGATGTCTCACCTTTGGACAGCATGAAAC 2383
617 rGlyYsSerProSerLysGlnIleMetArSerValIleSerValThr 634
2384 AGGTGGCTCTCCATCAAGCAACAGATGAGATCTGATTTCTGTACTT 2433
634 eArAlaLeuYsGluValGlyValaspSerSerleuThrAspThrArg 650
2434 CAGCTTTGAAGAGAGTTGGCGTGCAGAGATTAAGTAAATCCCGGAA 2483
651 ThrSerGluInlMetGlnlYrThrAsnAlaIleSerSerLysArg 667
2484 ACTTCAGAGAGATGCAAGAACCAACATGCTATTTCAGTAAGCGGAA 2533
667 uIleLeuArGArGleuasnGluInleuYsAlaGluInluspGluYs 684
2534 AATCTTCGACATTAATGAAAATCTTAAAGCTCAAGAGATGAAAAG 2583
684 lYmetGlnsnleuSerAspThrPheGluIleasnValHisGluAspAla 700
2584 GAATGCGAATCTCTGTGATCTTTGAGTAATGTTCATGAAAGATGCC 2633
701 LysGluHisGluLysGluLysSerValSerSerAspArgLysIleTyr 717
2634 AAGAGCATGAAAAGAAAAAATCACTTCACTCATGATGCAAGAAAGTGG 2683
717 uAlaGlyGluInleuValIleProLeuaspGluLeuThrleuAspThr 734
2684 GCGAGAGGTCAACTGTATTCCTGTGATGAGTACATACATATCAT 2733
734 ePheSerThrThrGluArGlnHisThrValGlyGluValIleLysLeu 750
2734 CCTCTCTACACTGAAAGACATACAGTGGGAGAAATTAATTAATAGT 2783
751 ProasnGlySerProArGArAlaTyrGlyLysSerProThrAspSer 767
2784 CCTAATGATCTCCAAAGAGAGCTGGGAAAGTCCGACAGATTCGT 2833
767 lLeuYsIleLeuGlyGluAlaGluLeuGluInleuInlThrGluLeu 784
2834 TCTAAATATCTTGAGAACTGAATCACTCAGACAGAACTATTAG 2883
784 lUasnThrThrIleArGSerGluIleSerProGluGlyLysIleYs 800
2884 AAATACACATTTAGAGTAGATTTCTCCGAAAGGGAAGAAAGTACAA 2933
801 ProleuIleThrGlyLysLysValGlnYsIleSerHisGluIleAs 817
2934 CCTTAATTTACTGGAAAAAAGTACAAATGATTTCCATGAAATAA 2983
817 nProSerAlaIleValAspSerProValGluThrLysSerProGluP 834
2984 CCCATCAGCTATGTGTATCTCTGTGAGACAAAAAATCCCGATTCA 3033
834 eGluAlaSerProGlnMetSerleuYsLeuGluInlYsInleuGlu 850
3034 GTGAGGCATCTCCACAGATGATGAACTGAACTGAAAGAAATTTAG 3083
851 ProAspArPleuGluThrGluIleLeuGluInlProSerGlyThr 867
3084 CCTGATGATTTGAAACAGAAATTTCTACAAAGGCAAGTGAACAA 3133
867 sAspGluSerleuProCysThrIleThrAspValTyrIleSerGlu 884
3134 AGATGAGAGCTTCCATCCATCTATTTACTGATGTGTGATTAGAGAA 3183
884 ySgluThrLysGluThrGlnSerAlaAspArgIleThrIleGlnGlu 900
3184 AAGAAACAAAGCAAACTAGTGGCAGATGAGATCACCATTGCGAAAT 3233
901 GluValSerGluAspGlyValSerSerThrValAspGlnleuSer 917
3234 GAAGTTCTGAAGATGAGATCGAGTACTGTGACCACTTAGTGCAT 3283
917 eHisIleGluProGlyThrAsnAspSerGlnHisSerLysCysAsp 934
3284 TCAATNAGAGCTCGAAGCAATGATTCACACATCTAAATAGTAGTAG 3333
934 sPlySerValGlnProGluProPhePheHisValValHisSerGlu 950
3334 ATAGCTGTGCACCGAAGCAATTTCCATTAAGGTGTCATTCGAA 3383
951 HisleuAsnleuValProGluValGlnSerValGlnYsSerProGlu 967
3384 CACTTAACCTTATGCTCAAGTTCATACATGATGCTGTACAGAGAA 3433
967 userPheAlaPheArGSerHisSerHisleuProProLysAsnLys 984
3434 ATCCTTTGCAATTCGATCTCAGCATTTTACACCAAAAAATTAACA 3483

```

```

984 ysaasSerleuLeuIleGlyLeuSerThrGlyLeuPheaspAlaasn 1000
|||||
3484 AGAATTCCTTGCTGATTCGACCTTTCACGTGCTGTTGATGCAACAC 3533
|||||
1001 ProLysMetLeuArgThrCysSerLeuProAspLeuSerLysLeuPhe 1017
|||||
3534 CCAAGAGATTAAAGACATGTTCACTCCAGATCTCCTCAAGCTGTTCAG 3583
|||||
1017 gThreuleMetAspValProThrValGlyAspValArgGlyAspAsnLeu 1034
|||||
3584 AACCTTATGATGTTCCACCGTAGGAGATGTTGCTCAAGACATCTTG 3633
|||||
1034 luileAspGluileLysAspGluAsnileLysGluGlyProSerAspSer 1050
|||||
3634 AAATAGATGAATTTAAAGATGAAAACATTAAGAAAGACCTTCTGATTC 3683
|||||
1051 GluAspIleValPheGluGluThrAspThrAspLeuGluLeuGlnAl 1067
|||||
3684 GAAGACATTTGTTGTAAGAACTGACACAGATTTACAGAGCTGAGGC 3733
|||||
1067 aserMetGluLeuLeuLeuArgGluGlnProGlyGluGlySerGlu 1084
|||||
3734 CTCGATGGAACAGTTACTTAGGAAACACCTGTAAGAAATACAGTCAAG 3783
|||||
1084 luGluGluSerValLeuLysasnSerAspValGluProThrAlaasnGly 1100
|||||
3784 AAGAAAGACGACGCTTGAAGAAACAGTGTGTGAGCCCACTGCAAAATGGG 3833
|||||
1101 ThrAspValAlaAspGluAspAsnProSerSerGluSerAlaLeuAs 1117
|||||
3834 ACAGATGTGGCGAGATGAAGATGACAAATCCAGTAGTCAAAAGTGCCTGAA 3883
|||||
1117 nGluGluThrPheHisSerAspAsnSerAspGluIleAlaSerGluGly 1134
|||||
3884 CGAAGAAATGGCACCTCAGATTAACAGTGAATGCGAAATTCGTAAGTCA 3933
|||||
1134 luCysAspSerValPheAsnHisLeuGluGluLeuArgGluHisLeuGlu 1150
|||||
3934 ATTCGGATGATGCTCTTTTACCAATTAAGAGAACGACATTCATCTGAG 3983
|||||
1151 GlnGluMetGlyPheGluLysPheGluValIleGlyLysIleLysAl 1167
|||||
3984 CAGGAAATGGCTTGAAGAAATCTTTCAGAGTTATAGAAATAAGAGC 4033
|||||
1167 alileHisGluAspGluAspGluAsnileGluileCysSerLysIleValG 1184
|||||
4034 TATTCATGACAGATGAAGATGAATATGAAATTTGTTCAAAATACTTC 4083
|||||
1184 lnaAsnileLeuGluLysGlnHisGlnHisLeuTyraIleLysIleLeuHis 1200
|||||
4084 AAAATATTTTGGGAAATGACATCCATCTTTATGCCAAGATCTTCAT 4133
|||||
1201 LeuValMetAlaAspGlyAlaIleTyraGlnLysAspAsnGlu 1214
|||||
4134 TTAGTCATGCGAGATGAGACCTTACCAAGACATATATATGAA 4175
|||||

```

seq_name: /net/abss06/STDS1/gc9data/hold-geneseq/geneseqn-emb1/NA2001A.DAT-AA158367

seq_documentation_block:

ID AA158367 standard; cDNA: 5448 BP.

AA158367:

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 570.

Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-drager Syndrome; chemotactic;
 chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

```

KW Leukaemia; ss.
XX
XX Homo sapiens.
OS
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX
XX 25-APR-2000; 2000US-0552317.
XX
XX 09-JUL-2000; 2000US-0598042.
XX
XX 19-JUL-2000; 2000US-0620312.
XX
XX 03-AUG-2000; 2000US-0653450.
XX
XX 14-SEP-2000; 2000US-0662191.
XX
XX 19-OCT-2000; 2000US-0693036.
XX
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HISEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AU, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX P-PSDB; AAM39211.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 570; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with noctropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localized neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: immune system suppression,
XX activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
XX Sequence 5448 BP; 1862 A; 932 C; 1168 G; 1486 T; 0 other;
XX

```

alignment_scores:

Quality: 6237.00

Length: 1214

Percent Similarity: 99.918

Gaps: 0

Percent Identity: 99.918

Percent Identity: 99.918

alignment_block:

us-09-783-320-4 x AA158367 ..

Align seg 1/1 to: AA158367 from: 1 to: 5448

```

1 MetGluLysTyraValArgLeuGlnLysIleGlyGluGlySerPheGlyLys 17
|||||
534 ATGGAAGAAGTATGTACATACAGAAAGATGCGAAGAGTTCATTGGAAA 583
|||||
17 SAAlileuValLysSerThrGluAspGlyArgGlnTyraIleLysG 34
|||||
584 AGCCATTCCTGTAAATCTACAGAAAGATGGCAGACGATGTATCAAGG 633
|||||
34 luileAsnileSerArgMetSerSerLysGluArgGluLysArgArg 50
|||||

```


634 AATTAACATCTCAAGATGTCACGTAAGAAAGAAAGAAATCAAGCAGA 683
51 GluValAlaValLeuLeuAlaSerMetLysHisProAsnIleValGlnTyr 67
684 GAACTTCACGATCTATTGCAACAAACATGAAGCATCCAAATATGTCAGTATAG 733
67 GGUUSePheGluGluGlnGlySerLeuTyrIleValMetAspTyrCysG 84
734 AGAATCTTTCAGAAAAATGGCTCTCTCTACATAGTAATGATTAACGTG 783
84 IugIlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
784 AGGAGGGGACTGTGTTTAACGCAATAATCTCAGAAAGGCGTTTGTT 833
101 GlnGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeu 117
834 CAAGAGATCGATTTGGACTGTGTGACATGACATATGTTTGCCCTCAA 883
117 SHISValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsn 134
884 ACATGTACATGATAGAAAAATCTCTCATCGACACATTAAATCTCAGAA 933
134 IephLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
934 TATTTTAACTAAAGATGAACAGTACAACTTGAGATTTTGGAATGCT 983
151 ArgValLeuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr 167
984 AGAGTCTTATAGTACTGATAGAGCTGGCTCGAACTTGATAGGGACCC 1033
167 GYTYTLLeuSerProGluIleCysGluAsnLysProTyrAsnAsnLys 184
1034 ATACTACTGTGCACCTGAAATCTGTGAACAAACCTTACAAATAAATAAA 1083
184 eTAspIleTPrAlaLeuGlyCysValLeuTyrGlnLeuCysThrLeuLys 200
1084 GAGACATTTGGGCTCTGGGTGTCTCTTATGAGCTGTGTACACTAAA 1133
201 HisAlaPheGlnAlaGlySerMetLysAsnLeuValLeuLysIleIle 217
1134 CATGCTTTTGAAGCTGCAGTATGAAAAACCTGCTACTGAATATATATC 1183
217 rGlySerPheProValSerLeuHisTyrSerTyrAspLeuArgSer 234
1184 TGGATCTTTCCACCTGTCTCTTTCANTATTCCTATGATCTCGCAGTT 1233
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
1234 TGGTGTCTCAGTTATTTAAAGAAATCTAGGATAGACATCAGTCAC 1283
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGlnLysPheLeu 267
1284 TCCATTTTGGAGAAAGTTTATAGCCAAAGCATGAAAGTTTCTCTC 1333
267 rProGlnLeuIleAlaGlnLeuPheCysLeuLysThrPheSerLysPheG 284
1334 TCTCTAGCTTATTTGCAAGAAATTTTGTAAACATTTTCGAAGTTG 1383
284 LysSerGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
1384 GATCAAGGCTATACAGCTAAAGAACAGCTTCAGGACAAACATTCGATT 1433
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGly 317
1434 TCTGTATGCTGCTCAGAAAAATTAACAAAGCCTGCCGTAAATATGAAAT 1483
317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGlnLysLysP 334
1484 ACCTTTGACATTAAGAAATATGAGATTAATAATTAACAGAAAAAAGAAC 1533
334 rGlnGlnLysHisLysGlnAlaHisGlnThrProGlnLysArgValAsn 350
1534 CACTGCAGAAAAACATTAACAGGCCCATCAATCTCAGAGAAAGAGTGAAAT 1583

351 ThrGlyGlnGluArgArgLysIleSerGlnGlnAlaAlaArgLysArg 367
1584 ACTGGGAGAAAGAGAAATATCTGAGAGACAGCAAGAAAGAAAG 1633
367 GLeuGlnPheIleGlnLysGlnLysLysGlnLysAspGlnIleIleSer 384
1634 GCTGGAATTTATTAAGAAAAAGAAACAAACAAAGATCGATTTATGCTT 1683
384 euMetLysAlaGlnGlnMetLysArgGlnGlnLysGlnLysGlnLysArg 400
1684 TAAATGAAGGCTGAAACAAATTAAGAAAGCAAGAAAGAAAGGCTGAAGA 1733
401 IleAsnArgAlaArgGlnGlnIlyTyrParAsnValLeuSerAlaGly 417
1734 ATAAATAGGCGCAGAGAAACAGATGAGAAATGTCTAAATGCTGAGTG 1783
417 YSerGlyValValLysAlaProPheLeuGlySerGlyGlyThrIleAla 434
1784 AAGTGTGAAGTAAAGGCTCTTTCTGGCAGCTGAGGAGCTATAGCTC 1833
434 rSerSerPheSerSerArgGlyGlnTyrGlnHisTyrHisAlaIlePhe 450
1834 CATCATCTTTTCTTCGAGAGACATGATGATACATTAACATCCATTTT 1883
451 AspGlnMetGlnGlnArgAlaGlnLysAspGlnAlaLysTPrLysArg 467
1884 GACCAATGCAAGCAACAAAGAGCAGAGATTAATGAACCTAAATGGAAG 1933
467 gGluIleTyrGlnArgGlyLeuProGluArgGlnLysGlyGlnLeuAla 484
1934 AGAAATATATGTGAGAGCTCTCCAGAAAGCAAAAAAGGCGAGCTAGTG 1983
484 aGlnUArgAlaLysGlnValGlnGlnPheGlnArgLysArgGlnAla 500
1984 TAGAAAGACTTAACAAATTAAGAAAGTCTCGACAGCAAAAGGGAAGCT 2033
501 MetGlnAsnLysAlaArgAlaGlnGlnHisMetValTyrLeuAlaArg 517
2034 ATCGAATTAAGCTGAGCCAGCAAGACATATGTTATCTGCGCAAGACT 2083
517 uArgGlnIleArgLeuGlnAsnPheAsnGlnUArgGlnGlnIleAlaAla 534
2084 GAGGCAATTAAGCTACAGAAATTTCAATGAGCGCAACAGATTAAAGCCA 2133
534 YsLeuArgGlyGlnLysLysGlnAlaAsnHisSerGlnGlyGlnGly 550
2134 AACCTGTGCTGAAAAAGAAAGAACTAATCTTGAAGGACAAAGAG 2183
551 SerGlnGlnAlaAspMetArgArgLysLysIleGlnSerLeuLysAla 567
2184 AGTGAAGAGGCTGACATGAGCGCAAAAAAATCGAAATCAGTGAAGGCCA 2233
567 sAlaAsnAlaArgAlaAlaValLeuLysGlnGlnLeuGlnUArgLysArg 584
2234 TGCAAATGCAAGCTGCTGTCTGTAAAGAAACATCGAACGAAAGGAA 2283
584 YsGlnAlaTyrGlnUArgGlnLysValTyrGlnGlnHisLeuValAla 600
2284 AGGAGCTTATGAGAGAAAAAAGTGTGGGAAGACATTTGCTGCT 2333
601 LysGlyValLysSerSerAspValSerProPheLeuGlnHisGln 617
2334 AAGAGATTAAAGTTCTGATGTTTCTCCACCTTTGGGACAGCATTAAC 2383
617 rGlyGlySerProSerLysGlnGlnMetArgSerValIleSerValThr 634
2384 AGGTGGCTCTCCATCAAGCAACAGATGAGATCTGTTATTTCTGTAAC 2433
634 eAlaIleLeuLysGlnValGlyValAspSerSerLeuThrArgThrArgGlu 650
2434 CAGCTTGAAGAGAGTTGGGTGACAGTATTTAACTGATACCGCGGAA 2483

651 ThrSergLumMetGlnLysThrAsnAsnAlaIleSerserLysArgI 667
 2484 ACTTGAGAGAGATGCAAAAGCCAAACATGCTATTTCAAGTAAGCGAGA 2533
 667 uilleuArGArgLeuAsnGluAsnLeuLysAlaGlnGluAspGluLysG 684
 2534 AATGCTGATGATTAATGAAAAATCTTAAGCTCAAGAGATGAAAAAG 2583
 684 lYMeGlnAsnLeuSerserPheGluIleAsnValIHisGluAspAla 700
 2584 GAAACGAGATCTCTGATACCTTTTGAGATTAATGTTCAATGAAGATGCC 2633
 701 LysGluHisGluLysGluLysSerValSerserAspArgLysLysTSPGI 717
 2634 AAGAGCATGAAAAAAGAAAAATCAGTTTCATCTGATCGCAAGAGATGGGA 2683
 717 uAlaGluGlnLeuValIleProLeuAspGluLeuThrLeuAspThs 734
 2684 GCGAGAGGTCAACTTGATGATTCCTGAGATGAGTTAACTAGATACAT 2733
 734 ePheSerThrThrGluArgHisThrValGluValIleLysLeuGly 750
 2734 CCTTCTCTCACTGAAAGACATACAGTGGAGAGAAAGTTAAATTAAGGT 2783
 751 ProAsnGlySerProArgArgAlaTSPGI LysSerProThrAspSerVa 767
 2784 CCTAATGATCTCCAAAGAGAGCTGGGGGAAAGTCCGACAGATCTGT 2833
 767 lLeuLysIleLeuGluGluAlaGluLeuGlnLeuGlnThrGluLeuLeuG 784
 2834 TCTAAAGATCTTGAGAGAGCTGAACCTCACTCAGACAGAACTATTAG 2883
 784 lAsnThrThrIleArgSerGluLysSerProGluGluGluLysTyrLys 800
 2884 AAAATACAACTATTGAAAGTGAATTTCTCCGAAAGGGAAGAAAGTCAAA 2933
 801 ProLeuIleThrGluLysLysValGlnCysIleSerHisGluIleAs 817
 2934 CCTTATATCTGAGAAAGAAAGATGATATTTTCACATGAAATTA 2983
 817 nProSerAlaIleValAspSerProValGluThrLysSerProGluPheS 834
 2984 CCTATGATGATTTGATGATTCCTCGTGGAGACAAAAAGTCCGAGTTCA 3033
 834 eGluAlaSerProGlnMetSerLeuLysLeuGluGluLysAsnLeuGlu 850
 3034 GTGAGCATCTCCACAGATGTCATTGAAACGTGAAGAAATTTAGAAAGAA 3083
 851 ProAspAspLeuGluThrGluIleLeuGlnGluProSerGluThrAsnL 867
 3084 CCTGATGATTTGAAAGCAAAATTTCTACAAAGAGCCAAAGTGACAAACAA 3133
 867 sAspGluSerLeuProCysThrIleThrAspValTSPGI eSergLumL 884
 3134 AGATGAGAGCTTGCCATGCACATTAAGTGTGATTAAGTAGAGAAA 3183
 884 ySGluThrLysGluThrGlnSerIleAspArgIleThrIleGlnGluAsn 900
 3184 AAGAAACAAAGGAAACTCAGTCGCGAGATAGATCACCATTCCAGAAAT 3233
 901 GluValSerGluAspGluValIleSerSerThrValAspGluLeuSerAsp 917
 3234 GAAGTTTCTGAAAGATGAGATCTCGAGTACTGAGCAACTTATGTCAT 3283
 917 eHisIleGluProGluThrAsnAspSerGlnHisSerLysCysAspAla 934
 3284 TCATATAGAGCTGGAGAACCAATGATTCACACACTCAAAATGTCATATAG 3333
 934 sPLysSerValGlnProGluProPhePheHisLysValIleHisSerGlu 950
 3334 ATTAAGCTGTGCAACCGAGAACCAATTTTCCATTAAGGTGTCATTCGAA 3383
 951 HisLeuAsnLeuValProGluValGlnSerValGlnCysSerProGlu 967

3384 CACTTGAAGACTTAGTCCCTCAAGTTCAATCAGTTGATGTTCCCGAGAAGA 3433
 967 userPheAlaPheArgSerHisSerHisLeuProProLysAsnLysAsnL 984
 3434 ATCTTTGCAATTCGATTCATCTCAGTCGATTTACCAACAAAAATTAACAA 3483
 984 yAsnSerLeuLeuIleGluLysSerThrGluLeuPheAspAlaAsnAsn 1000
 3484 AGAATTCCTTGATGATGAGCTTCAACTGAGTCTGTTGATGATCAAAACAC 3533
 1001 ProLysMetLeuArgThrCysSerLeuProAspLeuSerLysLeuPheAr 1017
 3534 CCAAGATGTTAAGACATGTTCTCCTCCAGATCTCTCAAGCTTTTCAG 3583
 1017 gThrLeuMetAspValProThrValGluLysPylArgGlnAspAsnLeuG 1034
 3584 AACCTTATGAGATGTTCCACCTGAGAGATGTCGTCAAGACAAATCTTG 3633
 1034 LuIleAspGluIleLysAspGluAsnIleLysGluGluProSerAspSer 1050
 3634 AAATAGATGAATTAAGATGAAGAAACATTAAGAAAGAGACCTTGATTC 3683
 1051 GluAspIleValPheGluGluThrAspThrAspLeuGlnGluLeuGlnAl 1067
 3684 GAAGACATTTGTTGTTGAAGAACTGACACAGATTTTACAAGAGCTGCAGC 3733
 1067 aSerMetGluGlnLeuLeuArgGluGlnProGluGluGluLysTyrSerGlu 1084
 3734 CTCGATGAACAGTTACTTAGGGAACACCTGAGTGAAGATACAGTGAAG 3783
 1084 LuGluGluSerValIleLysAsnSerAspValGluProThrAlaAsnGly 1100
 3784 AAGAAGATCAGTCTTGAAGAACAGATGATGAGGCAACCTCAATATGG 3833
 1101 ThrAspValAlaAspGluAspAspAsnProSerSerGluSerAlaLeuAs 1117
 3834 ACAATGTGCGAGTGAAGATGACAAATCCAGCAGTGAAGAGTCCCTGAA 3883
 1117 nGluGluThrHisSerAspAsnSerAspGluIleAlaSerGluCysG 1134
 3884 CGAAGATGCGACCTCAGATTAACAGTATGATGTAATTTGCTACTGAAATG 3933
 1134 lCysAspSerValPheAsnHisLeuGluGluLeuArgLeuHisLeuGlu 1150
 3934 AATCGATAGTGTCTTAAACCATTTAGAGAACTGAGACTTCATCTGAG 3983
 1151 GlnGluMetGlyPheGluLysPhePheGluValIleLysGluLysAl 1167
 3984 CAGCAAAATGGGCTTTGAAAATTTCTTTGAGGTTATGAGAAAAATTAAGC 4033
 1167 aIleHisGluAspGluAspGluAsnIleGluIleLysSerLysIleValG 1184
 4034 TATTCATGAAGATGAAGATGAAGAAATTTGAAATTTGTTCAAAAATAGTTC 4083
 1184 lAsnIleLeuGluLysGlnHisGlnHisLeuTyrAlaLysIleLeuHis 1200
 4084 AAAATATTGTTGGAAATGAAGACATCAGCATCTTATGCCCAAGTTTTCAT 4133
 1201 LeuValMetAlaAspGluAlaTyrGlnGluAspAsnAspGlu 1214
 4134 TTAGTCAATGCGAGTGGAGCTTCCAAAGAGATTAAGATGA 4175

seq_name: /nct/abss06/SDSI/9cgcdata/hold- geneseq/geneseg- emb1/NA2001A.DAT:AA158366
 seq_documentation_block:
 ID AA158366 standard: cDNA: 5532 BP.
 XX
 AC AA158366;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide seq ID NO 569.

XX Human: nocrotropic; immunosuppressant; cyostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW Leukemia; ss.
 XX Homo sapiens.
 XX MO200153312-A1.
 XX 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US34263.
 XX 21-JAN-2000; 2000US-0488725.
 XX 25-APR-2000; 2000US-0553317.
 XX 09-JUL-2000; 2000US-0596042.
 XX 19-JUL-2000; 2000US-0620312.
 XX 03-AUG-2000; 2000US-0653450.
 XX 14-SEP-2000; 2000US-0662191.
 XX 19-OCT-2000; 2000US-0693036.
 XX 29-NOV-2000; 2000US-0727344.
 XX (HYSEQ-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang Z, Wehrman T, Xu C, Xue AU, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 XX P-PSDB: AAM39210.
 DR Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 XX Claim 1; SEQ ID NO 569; 10078bp; English.
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM36642-AA44213) with nocrotropic,
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies, and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 SO Sequence 5532 BP: 1892 A; 949 C; 1195 G; 1496 T; 0 other;

alignment_scores: Length: 1242
 Ratio: 5.122 Gaps: 1
 Percent Similarity: 97.665 Percent Identity: 97.665

alignment_block:
 US-09-783-320-4 x AA158366 ..

Align seg 1/1 to: AA158366 from: 1 to: 5532

1 MetLuLysTyrValArgLeuGlnLysIleGlyGluGlySerPheGlyLys 17
 |||||||
 534 ATGGAGAGTATGTTAGCTACACAGAGATTGGAGAGAGTTTCATTGGAAA 583

17 GAlAlIleLeuValIysSerThrGluAspGlyArgIntYrValIleLysG 34
 |||||||
 584 AGCATTCTTGTATTAACCTACACAGATGCGACAGTATGTTATACAGG 633
 |||||||
 34 LuIleAsnIleSerArgMetSerSerLysGluArgGluGlnSerParg 50
 |||||||
 634 AAATTAACTCTCAGAAATGCTCCAGTAAAGAAAGAAATCAACAGSAGA 683
 |||||||
 51 GluValAlaValLeuAlaAsnMetLysHisProAsnIleValGlnTyrAr 67
 |||||||
 684 GAAGTTGCAGTATGCGAACAATGACATCCAAATATGTTGTCATATAG 733
 |||||||
 67 gLusErPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
 |||||||
 734 AGAATCATTTGAAAGAAATGCGTCTCTCATATGATGATGATGATGCG 783
 |||||||
 84 LuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
 |||||||
 784 AGCGAGGGGATCTGTTTAAGCGAATAAATGCTCAAGAAAGCGTTTGT 833
 |||||||
 101 GlnGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeuLys 117
 |||||||
 834 CAAGAGATCAGATTTGGACTGCTGTGTACAGATATGTTGGCCCTGAA 883
 |||||||
 117 sHtValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsnI 134
 |||||||
 884 ACATGTACATGATGAAAAAATTTCTCATCGAGCATTAATCTCAGAAC 933
 |||||||
 134 IepheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
 |||||||
 934 TATTTTACCTAAAGATGAGAACATACACTGTGAGATTTGGAAATGCT 983
 |||||||
 151 ArgValLeuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr 167
 |||||||
 984 AGAGTCTTAATAGTACTGTAGACCTGGCTGCAACTTGCTAGGACCCG 1033
 |||||||
 167 cTyTyrLeuSerProGluIleCysGluAsnLysProGlyAspAsnLysS 184
 |||||||
 1034 ATACTACTGTCACCTGAAATCTGGAAGAACAACTTAAATATATATAA 1083
 |||||||
 184 eArAspIleThrAlaLeuGlyCysValLeuTyrGlnLeuCysThrLysL 200
 |||||||
 1084 GTACACTTTGGGCTCTGGGGTGTGCTTATATGACCTGTGATACCTTAA 1133
 |||||||
 201 HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleIleSe 217
 |||||||
 1134 CATGCTTTTGAGCTGGCGAGTATGAAAAAAGCTGTACTGAAATATATATC 1183
 |||||||
 217 rGlySerPheProProValSerLeuHisTyrSerTyrAspLeuArgSerL 234
 |||||||
 1184 TGGATCTTTTCCACCTGTGCTTTCATATATCTATGATCTCCGACGTT 1233
 |||||||
 234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
 |||||||
 1234 TGGGTCTCGATTATTAAAGAAATCCAGGATAGACCATTCAGTCAAC 1283
 |||||||
 251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGlyLysPheLeuSe 267
 |||||||
 1284 TCCATATTGGAGAAAGTTTATAGCAAAAGCATTAAGAAAGTTTCTCTC 1333
 |||||||
 267 rProGlnLeuIleAlaGluGluPheCysLeuLysThrPheSerLysPheG 284
 |||||||
 1334 TCTCTATATGCTCTCGAAGAAATTAACAAAGCCGCTCAAAATATAGGAT 1483
 |||||||
 284 LysSerGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
 |||||||
 1384 GATCAGACGCTATACCAAGCTAAAGACAGCTTACAGCAACAACTCATTT 1433
 |||||||
 301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGlyI 317
 |||||||
 1434 TCTGTATATGCTCTCGAAGAAATTAACAAAGCCGCTCAAAATATAGGAT 1483
 |||||||
 317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGluLysLysP 334

```

1484 ACCCTTACCATAATAGAAATATGAGATATAAAATTCACGAAAGAAC 1533
334 rleuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
1534 CACTGCAAAACATTAACAGAGCCCATCAAACTCCAGAGAGAGAGTCAT 1583
351 ThrGlyGluGluArgArgLysIleSerGluGluAlaIleArgLysArg 367
1584 ACTGGGAGAAAGAGAGAAATATCTGAGGAAACACAGCAAGAAAGAG 1633
367 GLeuGluPheIleGluLysGluLysGlnLysAspGlnIleIleSer 384
1634 GCTGGAATTTATGAAAAGAAAAGAAACAAAGAGATCAATATTATG 1683
384 euMetLysAlaGluGluMetLysArgGlnGluLysGluArgLeuGlu 400
1684 TATGTAGAGCTGAACAAATGAAAGGCAAGAAAGAAAGGTTGGAAGA 1733
401 IleAsnArgAlaArgGluGlnGlyTrpArgAsnValLeuSerAlaGly 417
1734 ATTAATAGGCGCAGGAAACAAAGATGAGAAATGCTTAAGTCTGTGG 1783
417 ySerGlyValLysAlaProPheLeuGlySerGlyGlyThrIleAla 434
1784 AAGTGTAGAGTAAAGGCTCTTTCGCGCAGTGGAGGACATATAGCT 1833
434 rSerSerPheSerSerArgGlyGlnTrpGluHisIleAlaIlePhe 450
1834 CATCATCTTTTCTCTCGAGACAGATATGAACATTACCATTCCTTT 1883
451 AspGlnMetGlnGlnIleArgAlaGluAspAsnGluAlaLysTrpLys 467
1884 GACCAATGCGAGCAACAAAGAGAGAAATATGAACCTTAATGGAAAG 1933
467 GGlLysLeuTrpGlyArgGlyLeuProGluArgIleLysGlyGluLeu 484
1934 AGAAATATATGTGTCAGGCTCTCCAGAAAGGCAAAAGGCGACTGAC 1983
484 aIGluArgAlaLysGlnValGluGluPheLeuGlnArgLysArgGlu 500
1984 TAGAAAGAGCTTAACAAAGTAGAAGATTCCTGCGAGCAAAAGCGA 2033
501 MetGlnAsnLysAlaArgAlaGluGlyHisMet..... 511
2034 ATGCAAGATTAAGCTCGAGCGAAGACATATGGAAATCCTGCAAA 2083
511 ..... 511
2084 GGCACATATGTATGAGGCGAGGCCAGCTTCAAGAGAGAGGAGCA 2133
512 .....ValTrpLeuAlaArgLeuArgIleArgLeu 522
2134 GAACAAAGAGAGAGAGCTTATCTGCAAGAGCTGAGCAAAATAGAC 2183
523 GlnAsnProAsnGluArgGlnGlnIleLysAlaLysLeuArgGlyGly 539
2184 CAGATTTTCATGAGCGCCAAACATTAACCCAACTCTGCTGTAAAA 2233
539 sArgGluAlaAsnHisSerGluGlyGlnGlySerGluGluAlaAsp 556
2234 GAAAGAGCTATATCATCTCTGAAGACAGAGAGGAGAGGCTGTACA 2283
556 eArgArgLysLysIleGluSerLeuLysAlaHisAlaAsnAlaArg 572
2284 TGAGCGCGCAAAATTCGAATCACTGAAGGCCCATGCAAAATGCA 2333
573 AlaValLeuLysGluGlnLeuGluArgLysArgLysGluAlaTrpGlu 589
2334 GCTGTACTTAAAAAGAACACTAGAACGAAGAGAGAGGCTTATGTAG 2383
589 gLysLysLysValTrpGluGluHisLeuValAlaLysGlyValLys 606
1 .....
2384 AGAAAAAAGTGTGGAGAGCATTGGTGCCTAAAGAGCTTAAGACT 2433
606 eArgValSerProProLeuGlyGlnHisGluThrGlyLysProSer 622
2434 CTGATGTTTCTCCACCTTGGGACAGCATGAAGAGGTGGCTGTCA 2483
623 LysGlnGlnMetArgSerValIleSerValThrSerAlaLeuLysGlu 639
2484 AAGCAACAGATGAGATCTGTTATTTCTGTAACCTGACGTTTGAAGA 2533
639 IGLValAspSerSerLeuThrAspThrArgGluTrpTrpSerGluMet 656
2534 TGGCTGTGACAGTAGTTTAACCTGATCCGGGAACTTCAGAGAGAT 2583
656 IuLysThrAsnAlaIleSerSerLysArgGluIleLeuArgArgLeu 672
2584 AAAAGACCAACATGCTATTTCAGTAAAGCGAAGAAATACCTTGAT 2633
673 AsnGluAsnLeuLysAlaGlnGluAspGluLysGlyMetGlnAsnLeu 689
2634 AATGAAATCTTAAGCTCAAGAGATGAAAAAGAAAGCAAGAAATCT 2683
689 rAspThrPheGluIleAsnValHisGluAspAlaLysGluHisGly 706
706 IuLysSerValSerSerAspArgLysLysTrpGluAlaGlyGlnLeu 722
2734 AAAATACGTTTCATCTGATCGCAAGAGTGGAGCGACAGAGTCACT 2783
723 ValIleProLeuAspGluLeuThrLeuAspThrSerPheSerThrG 739
2784 GTGATCTCTGATGATGATTAACACTAGATACCTCTTCTACAACTG 2833
739 uArgHisThrValGlyGluValIleLysLeuGluProAsnGlySerPro 756
2834 AAGCATACAGTGGAGAACTTATTAATTAGTCTTAATGATGATCCAA 2883
756 rArgAlaIleTrpGlyLysSerProThrAspSerValLeuLysIleLeu 772
2884 GAAGAGCTGTGGGAAAGTCCGACAGATTCCTGTTTAAAGTACTGGA 2933
773 GluAlaGluLeuGlnLeuGlnThrGluLeuLeuGluAsnThrIleAr 789
2934 GAACCTGACTACAACTTCAGACAGAACTATGAATAACATATTAAG 2983
789 gSerGluIleSerProGluGlyLysLysTrpLysProLeuIleThrGly 806
2984 AAGTGAGATTTCTCCGAGAGGGAAGATTAACCCCTTAATTAAGT 3033
806 IuLysLysValGlnCysIleSerHisGluIleAsnProSerAlaIle 822
3034 AAAAAAAGTACATGATTTTACATGAATAAACCACATCACTATGCT 3083
823 AspSerProValGluTrpTrpLysSerProGluPheSerGluAlaSerPro 839
3084 GATTCCTCTGTGAGCAAAAGTCCCGAGTTCAAGAGACATCTCCACA 3133
839 nMetSerLeuLysLeuGlnGlyAsnLeuGluTrpAspAspLeuGlu 856
3134 GATGTCAATTAAGCTGAGAGAAATTTAGAGAACTGTGATTTGGAAA 3183
856 hrGluIleLeuGlnGluProSerGlyThrAsnLysAspGluSerLeu 872
3184 CAGAAATTCATCAAGAGCCAAAGTGAAACAAACAAATAGAGCTTGC 3233
873 CysThrIleThrAspValTrpIleSerGluGluTrpThrLysGluTh 889
3234 TGCACTATTAAGTATGTGATTAAGTGAAGAAAAAGAAAGAAAGAAC 3283
889 rGlnSerAlaAspArgIleThrIleGlnGluAsnGluValSerGluAsp 906
3284 TCACTGCGCAGATAGATCACTCACTCAGGAAATGAAGTTCTGAAG 3333

```

906 lvalserSerThrValaspGlnleuSeraspIleHisIleGluProGly 922
 3334 GAGTCTCGAGTACTGTGACCAACTGATGACATTCATATAGAGCCGGA 3383
 923 ThrAspSerSerGlnHisSerIleCysAspValAspLysSerValGlnPr 939
 3384 ACCAATGATTCACAGACTCTAAATGTGATGATGATGATGATGATGATG 3433
 939 ogIuprophepHeHisValValHisSerGlnHisIleuasnleuValP 956
 3434 GGAACCAATTTTCATAGGTGCTTCATTCGAACTGAACTGAACTTACTGC 3483
 956 roGlnValGlnSerValGlnCysSerProGlnGluSerPheAlaPheArg 972
 3484 CTCAAGTTCATCATGATTCAGTTCACAGAGAAATCTTTCGATTCGA 3533
 973 SerHisSerHisleuProProllysasnLysasnLysSerleuLeuI 989
 3534 TCTCAGCTGCGATTACCCAAAATAAATAAACAAGAAATCTTTCGCTGAT 3583
 989 eGlyLeuSerThrGlyLeuPheaspAlaAsnAsnProLysMetLeuArgT 1006
 3584 TGGACTTTCAACGCTGCTTTTCATGCAAAACCAACCAAGATGTTAGGA 3633
 1006 hrCysSerLeuProaspLeuSerLysleuPheArgThrIleuMetAspVal 1022
 3634 CATGTCACCTCCAGATCTCTCAAAAGCTGTCAGAAACCTTATGATGTT 3683
 1023 ProThrValGlyAspValArgGlnAspAsnLeuGlnIleAspGluIleL 1039
 3684 CCCACCGTAGAGATGTTCTCAAGCAAACTTGAATGAGAAATTA 3733
 1039 sasPglAsnIleLysGlnGlyProSerAspSerGlnAspIleValPheG 1056
 3734 AGATGAAACATTAAGAGAGACCTTCTGATCTGAAAGCATGTCGCTTGG 3783
 1056 lgluThrAspThrAspLeuGlnGlnleuGlnIleAsnSerGlnGlnleu 1072
 3784 AAGAACTGCACAGATTTACAAAGAGCTGCAGGCTCGATGGAACGTTA 3833
 1073 leuArgGlnGlnProGlnGlnIleuThrSerGlnGlnGlnGlnIleuVal 1089
 3834 CTTAGGAAACACCTGCTGAGAGATTAAGTGAAGAAAGACAGCATCTTT 3883
 1089 ulysAsnSerAspValGluProThrAlaAsnGlyThrAspValAlaAspG 1106
 3884 GAAGAAACAGTGTGTGAGCCAACTGCAAAAGGACAGATGTGCGAGATG 3933
 1106 lAspAspAsnProSerSerSerGlnSerAlaIleuAsnGlnIleuThrP 1122
 3934 AAGATGACAAATCCAGAGTGAAGAGTCCCTGAACGAAAGATGCGACTCA 3983
 1123 AspAsnSerAspGlyIleAlaSerGlnCysGlnCysAspSerValP 1139
 3984 GATTAACAGTGTGATGAAATGCTAGTGAATGTAATGCAATGATGATGCT 4033
 1139 eAsnHisIleuGlnGlnleuArgLeuHisIleuGlnGlnIleuMetGly 1156
 4034 TAACCAATTAAGAGAACTGAGACTTCATCTGAGCAGAGAAATGGCTTGG 4083
 1156 lIuysPhePheGlnValIleuGlnIleuValIleuValIleuValIleu 1172
 4084 AAAATTTCTTGGAGTTTATGAGAAATTAAGGCTATTCATGAGATGAA 4133
 1173 AspGlnAsnIleGlnIleuLysSerLysIleValGlnAsnIleuGlnL 1189
 4134 GATGAAATATTTGAAATTTCTTCAAAATAGTTCAAAATATTTTGGGAAA 4183
 1189 nGlnHisGlnHisIleuThrAlaLysIleuHisIleuValIleuValIleu 1206
 4184 TGAACATGACAGATCTTATGATGCCAAGATTTCTTCAATTTAGTCAGAGATG 4233

1206 lvalATyrGlnGlnIuAspAsnAspGln 1214
 4234 GAGCTTACCAAGAGATATATGATGAA 4259
 seq_name: /net/abs06/STDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA2001A.DAT:AA160152
 seq_documentation_block:
 ID AA160152 standard; cDNA; 5514 BP.
 AC AA160152;
 DT 22-OCT-2001 (first entry)
 XX Human polynucleotide SEQ ID NO 4141.
 DE Human;
 XX Human; nototropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukemia; ss.
 XX Homo sapiens.
 OS
 PN WO200153312-A1.
 XX 26-JUL-2001.
 PD
 PF 26-DEC-2000; 2000WO-US34263.
 XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 DR 09-JUL-2000; 2000US-0598042.
 DR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR P-PSDB: AAM40996.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS Claim 1; SEQ ID NO 4141; 10078bp; English.
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AA442213) with nototropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC activity/inhibin activity, chemotactic/chemokine activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX Sequence 5514 BP: 1500 A; 1191 C; 943 G; 1880 T; 0 other;

Wed May 15 14:14:00 2002

us-09-783-320-4.rng

Page 12

alignment scores: Length: 1247
 Quality: 6097.00
 Ratio: 5.047
 Gaps: 6
Percent similarity: 96.872 Percent identity: 96.552

alignment block:
us-09-783-320-4 x AA160152/rev ..

Align seg 1/1 to reverse of: AA160152 from: 1 to: 5514

```
1 MetGluTyrValArgLeuGlnLysIleGlyLugLysPheGly 17
|||||
4981 ATCGAAGATGTTAGCTACAGAAAGTTGGAAAGCTTCTGAAA 4932
17 sAlaIleuValLysSerThrGluAspGlyArgGlnTyrValIle 34
|||||
4931 AGCCATTCTGTAAATCTACAGAAAGTGGCAACATATGTTATCAGG 4882
34 IuIleAsnIleSerArgMetSerSerLysGluArgGluGluSerArg 50
|||||
4881 AATTAAATCTCAAGAAATGTCAGTAAAGAAAGAAATCAAGAGA 4832
51 GluValAlaValLeuAlaAsnMetLysHisProAsnIleValGlnTyr 67
|||||
4831 GAAGTTGGAGATATGGCAACATGACATCCAAATTTCTCCAGTATAG 4782
67 gLusSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCys 84
|||||
4781 AGAATCATTTGAAAGAAATGGCTCTCTACATAGTAAATGATTAAGT 4732
84 LuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValIlePhe 100
|||||
4731 AGGAGGGGATCTGTTAAAGCAATTAATGCTCGAAAGCGCTTTCTT 4682
101 GlnGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeu 117
|||||
4681 CAAGAGATCAATTTTGACCTGGTGTACAGATATGTTGGCCGAA 4632
117 sHisValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsn 134
|||||
4631 ACATGTACATGATGAAAAATTTCTTCAGACATTAATCTCAGAAC 4582
134 IePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
|||||
4581 TATTTTAACTAAAGATGAAACAGTATTAATCGAGATTTTGGAAATGCT 4532
151 ArgValIleAsnSerThrValGlnLeuAlaArgThrCysIleGlyThr 167
|||||
4531 AGAGTTCTTAATAGTACTGTAGACTGGCTCGAATTGCAATAGGAGCC 4482
167 oTyTyrLeuSerProGluIleCysGluAsnLysProTyrAsnAsnLys 184
|||||
4481 ATACTACTTGTCACTGAAATCTGTGAAACAAACCTTACATATATATAA 4432
184 eAspIleThrAlaLeuGlnCysValLeuTyrGlnLeuCysThrLeuLys 200
|||||
4431 GTGAATTTGGGCTGTGGGGGTGTCTTATAGAGCTGTACACTATAA 4382
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleIle 217
|||||
4381 CATGCTTTGGAAGTCGCAATGAAAAAAGCTGTACTGAAGATAAATTC 4332
217 rGlySerPheProValSerLeuHisTyrSerTyrAspLeuArgSer 234
|||||
4331 TGGATCTTTTCCACCTGTGTTCATTTATCTATGATCTCCGCGAGTT 4282
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
|||||
4281 TGGTGTCTCAGTTATTAAGAAATCTCTAGAGATACCATCAGCTAAC 4232
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGluLysPheLeu 267
|||||
4231 TCCATATTTGGAAGATTTTATACCAAGCATGAAAAAGTTTCTCTC 4182
```

```
267 rProGlnLeuIleAlaGlnLys...PheCysLeuLysThr...PheSerLys 282
|||||
4181 TCCACACTTATTTGCAAGAAATTTTGTCTTAAAAAATTTTTCAGAG 4132
283 PheGlySerGlnProIle...ProAlaLysArgProAlaSerGlyLys 298
|||||
4131 TTTGGATCAACAGCTTATACAGCTTAAAGACAGCTTTCAGCAAAA 4082
298 nSerIleSerValMetProAlaGlnLysIleThrLysProAlaAlaLys 315
|||||
4081 CTGATTTCTGTTATGCTCTCAGAAATTTACAAAGCCCTGCGTAAAT 4032
315 yrgIlyIleProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHis 331
|||||
4031 ATGGAATACCTTTAGCTATTAAGAAATATGAGATTAATAAATTACACGA 3982
332 LysLysProLeuGlnLysHisLysGlnAlaHisGlnThrProGluLysArg 348
|||||
3981 AAGAAACCCACTGCAAAACATTAACAGCCCATCAAACTCCAGAGAGAG 3932
348 gValAsnThrGlyGlnGluArgArgLysIleSerGluGluAlaAlaArg 365
|||||
3931 AGTGAATTCCTGGGAGAAAGAAAGAAATATCTGAGAAAGCAGCAAGAA 3882
365 ySarGArgLeuGluPheIleGlyLysGlnLysLysAspGlnIle 381
|||||
3881 AGACAGAGCTGGAATTTATGAAAAAGAAACAAAGCAATGATCAGATT 3832
382 IleSerLeuMetLysAlaGlnGlnMetLysArgGlnGluLysGluArg 398
|||||
3831 ATTAGTTTATGAAAGCTGTAACAAATGAAAAAGCAAGAAAGAAAGTT 3782
398 uGluArgIleAsnArgAlaArgGlnGlnIleTyrPheArgAsnValIleSer 415
|||||
3781 GGAAGAAATATATAGGCGCAGGAAACAGATGAGAAATGCTGCTAAGTG 3732
415 IeGlyLysArgGlyGluValLysAlaProPheLeuGlnLysSerGlyThr 431
|||||
3731 CTGGTGGAGAGCTGGAAGTAAAGCTCTTCTGCGAGAGTCCAGAGACT 3682
432 IleAlaProSerSerPheSerSerArgGlyGlnTyrGlnHisAla 448
|||||
3681 ATACCTCCATCATCTTTTCTTCTCGAGAGCAGTATGAACATTAACATGC 3632
448 aIlePheAspGlnMetGlnGlnIleArgAlaGluAspAsnGluAlaLys 465
|||||
3631 CATTTTGGACCAATGCAGCAACAAAGACAGCAATGATGAACCTAAAT 3582
465 rPlySarGluIleTyrGlyArgGlyLeuProGluArgGlnLysGlyGln 481
|||||
3581 GGAAAGAGAAATATATAGTGCAGAGCTTCCGAAAGGCAAAAAAGGCGAG 3532
482 LeuAlaValGluArgAlaLysGlnValGlnGluPheLeuGlnArgLysArg 498
|||||
3531 CTAGCTGTAGAAAGCTTAACAAGTGAAGAGCTTCTCAGCGAAAGAG 3482
498 gGluAlaMetGlnAsnLysAlaArgAlaGlnLysMet... 511
|||||
3481 GGAAGCTATGAGAAATTAAGCTCGAGCCGAAAGCATATGGAATCTTGC 3432
511 ..... 511
3431 AAAACCTGGAGCTATATGAGAGCGAGCCAGCTTCAAGAGAGGG 3382
512 .....ValTyrLeuAlaArgLeuArgGlnIle 520
3381 AAGCAAGAAACAAAGAGAGAGGTTTATCTCGCAAGCTCAGCGCAAT 3332
520 eArgLeuGlnAsnPheAsnGluArgGlnGlnIleLysAlaLysLysArg 537
|||||
3331 AAGACTACAGATTTCAATAGAGCGCCACACAGATTAAGCCAAACTTCGTG 3282
```

337	lyglulysglulualasnsergluglmglnlgllysergluglu	553
3281	GTGAAAGAAAGAGCTAATCTATTCTGTAGACACAAAGCAATGTGAAG	3232
554	Alaaspetararg Lyslysilegluserleuylalalhisalasn	570
3231	GCCTGACGTGAGCGCAAAAAAATCGAATCTAGAGGCCCATGCCAAATG	3182
570	laargalalalaleuylsglmglnleuarglysarplysgluale	586
3181	CACCTGCTGCTTACTAAAGAACACTGAAACCAAAAGAGAGAGAGGCT	3132
587	tyrgluarglulyslysaltrpoglunhlisleuvalalaysgluya	603
3131	TATAGAGAGAAAAAAGCTGTGGAAAGCAATYTGCGCTAAAGAGT	3082
603	lysserfaspvalserprobleugllyglnhlsglulthrsgllyls	620
3081	TAAAGTTCTGATGTTCTTCCACCTTGGACAGCATGAAACAGTGGCT	3032
620	erproserlysglnlmetargservalleservalthrseralaleu	636
3031	CTCCATCAAAACCAACAGATGAGTCTGTATTCTGTACTTCCAGCTTGG	2982
637	lysglulvalglvalaspserserleuthraspthrarglyuphrsergl	653
2981	AAAGAGTTGGCTGGACACTACTTTAACTGATACCGGAAACTTCAGA	2932
653	ucllmetglulysthrnsnspalaleserterlysarplyleleua	670
2931	ACAGATGCAAAAGACCAACATGCTATTCCAGTAAAGCAGAAATACCTC	2882
670	rgargleuasnrgluasnleuylalaglnlualspglulysgllymergln	686
2881	GTAGATTAAATGAATAATCTTAAGCTCAAGAAAGATCAAAAAGAAAGAG	2832
687	asnleuseraspthrphleglulileasnvalhlsglualsyalysglunl	703
2831	AATCTCTCTGATACTTGTGAAATAAATGTCATGAAGAAGATGCCAAAGACA	2782
703	sglulysglulysservalserasparglylystrpglualaglyg	720
2781	TGAAAAAGAAAAATCAGTTTCATCTGATCGCAAGAAAGTGGAGCGAGAG	2732
720	lyglnleuvalleproleuaspgluleuthrleuaspthrserpheser	736
2731	GTCAACTTGTGATCTCTGATGAGTAACTGATACATCTCTCTCT	2682
737	thrthgluarghisthrvalglylvalilleuylsleugllyproasncl	753
2681	ACAACTGAAAGACATACAGTGGAGAAAGTATTAATTAAGTCTTAATGG	2632
753	ysrserproargalatrpglylysserprothrhaspservalleyysi	770
2631	ATCTCCAAAGAGAGCGCTGGGGAAAAAGTCCGACAGATCTGTTCTTAAGA	2582
770	leleuylglulagluuengluuengluhthrghuleuengluuasnthr	786
2581	TACTTGGAGAACTGAACACTTCCAGTTCAGACAGAACTATTAAGAAATACA	2532
787	thrllleargserglulieserproglunlglulyslytyrlyserleuul	803
2531	ACTATTGAGGTAGATTCTCCCAAGGGGAAAAAGTACAAACCTTAAT	2482
803	ethrglyglululysvalcincyslleserhisglulileasnprosera	820
2481	TACTGGAGAAAAAGTACAAATGATTTTCACATGAAATAAACCCATCAG	2432
820	lailevalaspserserprovalglulthrlysserprogluluphesergluala	836
2431	CTATTGTTGATTTCTCCGTGTGAGACAAAAAGTCCGAGCTTCAGTGAAGCA	2382
837	serproglumetserleuylsleuglulysleuengluuengluuproaspas	853

201 HisAlapheGluAlaGlySerMetLysAsnLeuValLeuLysIleIleSe 21

201 HisAlaphegluaIaGlySermettLysasnLeuValLeulysileilese 21


```
|||||
4381 CATGCTTTTACGTGACGATGAAAAACCTGTACTGAGATTAATATC 4332
217 rgllyserPheProValSerLeuHisTyrSerTyrAspLeuArgSerL 234
|||||
4331 TGGATCTTTTCCACCTGTGCTTTCATATATCTATGATCTCCGCACTT 4282
234 euValSerGlnLeuPheLysArgAspProArgAspArgProSerValAsn 250
|||||
4281 TGGGTCTCAGATTATTTAAAGAAATCTTAGGGATAGACATCACTCAAC 4232
251 SerLeuGlnLysGlyPheLeuAlaLysArgLLeuLysPheLeuSe 267
|||||
4231 TCCATATTGGAGAAAGSTTTTATAGCCAAACGATTTGAAAGTTTCTCTC 4182
267 rPrGlnLeuTLeuAlaGlnLysLysPheCysLeuLysThr.PheSerLys 282
|||||
4181 TCCCTCAGCTATTTGACGAAAGAAATTTTGTCTTAAACATTTTTCGAAAG 4132
283 PheGlySerGlnProIle...ProAlaLysArgProAlaSerGlyGlnAs 298
|||||
4131 TTTGGATCACAAGCCTTATACAGCTTAAAGACCAAGCTTCAGGACAAAA 4082
298 nSerLieserValMetProAlaGlnLysLLeuThrLysProAlaAlaLysT 315
|||||
4081 CTCGATTTCTGTATGCTGCTGCTGCAAAATTTACAAAGCCTCGCGCTAAAT 4032
315 YrGlyLeuProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGln 331
|||||
4031 ATGGATATCCTTAGCATATTAAGAAATATGAGATTAATAAAATTTACCGAA 3982
332 LysLysProLeuGlnLysHisLysGlnAlaHisGlnThrProGlnLysAr 348
|||||
3981 AAGAAACCATCTGCACAAACATTAACAGCCATCAAACTCCACAGAGAGAG 3932
348 gValAsnThrGlyGlnLysArgLysLLeuSerGlnLysAlaAlaArgL 365
|||||
3931 AGGAAATCTGGCGAAGAAAGAGAAATAATCTGAGGAGACAGCAAGAA 3882
365 YsrArgLysGlnLysPheLLeuLysGlnLysLysGlnLysAspGlnLle 381
|||||
3881 AGGAAAGCTGGCAATTTATCTGAAAAAGAAAAACAAAGATCGAGATT 3832
382 LLeuSerLeuMetLysAlaGlnLysLysLysArgGlnLysGlnLysArgL 398
|||||
3831 ATTAGTTTATGATGAGCTGAAACAAATGAAAGCAAGAAAGGAAAGTT 3782
398 uGlnArgLLeuAsnArgAlaArgLysGlnLysLysLysLysLysLysL 415
|||||
3781 GGAAGAGATTAATATAGGCCAGGAGCAAGAGATGAGAAATGCTTAAGTG 3732
415 lAcGlySerGlyGlnValLysAlaProPheLeuGlySerGlyGlyThr 431
|||||
3731 CTGGTGGAGATGGTGAAGTAAAGGCTCTTCTGCGCAGTGGAGGACT 3682
432 lLeuAlaProSerSerPheSerSerArgLysGlnLysLysLysLysL 448
|||||
3681 ATACCTCATCATCTTTTCTCTCTCGAGACAGTATGAACATTTTCCATGC 3632
448 aAlaPheAspGlnMetGlnGlnLysArgLLeuAlaLysLysLysLysL 465
|||||
3631 CATTTTGCACCAATCAGACACAAAGACAGACAGATATATGATGATTAAT 3582
465 rPlySerArgLysLysLysLysLysLysLysLysLysLysLysLysL 481
|||||
3581 GGAAGAAAGAAATATATGTCGAGGCTCTCCAGAAAGGCAAAAGGCGAG 3532
482 LeuAlaValGlnArgAlaLysGlnValGlnLysPheLeuGlnArgLysAr 498
|||||
3531 CTAGCTGAGAAAGAGCTAAACAAGTAAAGAGATCTCTCAGCGAAACG 3482
498 gGlnAlaMetClnAsnLysAlaArgLLeuGlnLysMet..... 511
|||||

3481 GGAAGCTATGCAATTAAGCTCGAGCCGAGAGACATATGGAATCTCTGC 3432
511 ..... 511
3431 AAAACCTGGCAGCTATGTATGAGAGCAGGCCACAGCTCTTAAGAGGAGG 3382
512 .....ValTyrLeuAlaArgLLeuArgGlnL 520
3381 AAGCCAAAGAAACAAAGAGAGAGGTTTATCTGGCAAGCTCAGAGCAAAAT 3332
520 eArgLysGlnAsnPheAsnGlnArgGlnLysLysLysLysLysLysLys 537
3331 AAGACTCAGAAATTTCAATGAGCGCCACACAGATTAAGCCAAACTGCTG 3282
537 lYgLnLysGlnLysLysLysLysLysLysLysLysLysLysLysLysL 553
3281 GTGAAAGAAAGAAAGCTAATCATCTGAAAGGACAAAGAGAGAGTGAAGAG 3232
554 AlaAspMetArgArg.LysLysLLeuLysSerLeuLysAlaHisAlaAsn 570
|||||
3231 GCTGACATGAGCGCCAAAAAAATCGAATCTCGAAGAGGCCCAATG 3182
570 lArgAlaAlaValLeuLysGlnLysGlnLysArgLysArgLysGlnAla 586
3181 CACGTGCTGCTGTACTTAAAGAACACTAGAACGAAAGAGAGAGGCT 3132
587 TyrGlnArgGlnLysLysValTyrGlnLysLysLysLysLysLysLys 603
3131 TATGAGAGAGAAAAAAAGTGCGGAGAGACATTTGGCTGAAGAGACT 3082
603 lLysSerSerAspValSerProProLeuGlnLysHisGlnThrGlyGly 620
3081 TAGAGTTCTGATGTTTCTCCACCTTGGGACAGCATGAAGACAGTGCT 3032
620 ePrProSerLysGlnLysMetArgSerValLLeuSerValThrSerAlaLeu 636
3031 CTCATCAAGACACAGATGAGATCTGTATTCTGTGAATCTCAGCTTG 2982
637 LysGlnValGlyValAspSerSerLeuThrAspThrArgLysLysSerG 653
2981 AAGAGAGTTGGCTGACAGTGTATTAAGTATATCCCGGAAACTCAAG 2932
653 uGlnMetGlnLysThrAsnAsnAlaLLeuSerLysArgLysLysLysL 670
2931 AGAGATGCAAAAGACCACAAATGCTATTTCAGTAAAGCGAAGAAATCTT 2882
670 rArgLysGlnLysGlnLysLysLysLysLysLysLysLysLysLysL 686
2881 GTAGATTAAATGAATCTTAAAGCTCAAGAGATGAAAAAGAAAGCAG 2832
687 AsnLeuSerAspThrPheGlnLysLysLysLysLysLysLysLysLys 703
2831 AATCTCTGATPACTTTTGAATTAATGTCATGAAAGATGCAAGAGCA 2782
703 sGlnLysGlnLysSerValSerSerAspArgLysLysTyrGlnLysL 720
2781 TGAAGAAAGAAATCAAGTTTCACTGATCGCAAGAGAGGAGGAGAGAG 2732
720 lYgLnLeuValLLeuProLeuAspGlnLysLysLysLysLysLysLys 736
2731 GTCACTTGTGATCTCTCGATGAGATTAACCTGATTAACATCTCTCT 2682
737 ThrThrGlnArgHisThrValGlyGlnValLLeuLysGlnLysProAsn 753
2681 ACAACTGAAGACATACAGTGGAGAGAAATTAATTAATTAAGTCCATATG 2632
753 YSerProArgArgAlaTyrGlyLysSerProThrAspSerValLeuLys 770
2631 ATCTCCAGAGAGAGCTGGGGGAAAAAGTCCGACAGATTCGTGTAAAGA 2582
770 lLeuGlyGlnAlaGlnLeuGlnLysLysLysLysLysLysLysLysL 786
2581 TACTTGGAGAGCTGAAGTCAACTTCAAGCAAGACTATTAAGAAAAATACA 2532
```

787 ThrIleArgSerGluIleSerProGluGlyIleuSerTyrLysProIleuI 803
 2531 ACTATTAGAGTGAAGATTTCTCCGAGGGGAGGAGAACAAACCCCTTAAT 2482
 803 ethrGlyGlyLysValGlnCysIleSerHisGluIleAsnProSerA 820
 2481 TACTGAGAGAAAAAGATCAATGATTTTCATGAAATAAACCATCAG 2432
 820 IatIleValAspSerProValGluThrLysSerProGluPheSerGluAlA 836
 2431 CTAATGTGATTCTCCTGTTGAGACAAAAGTCCCGAGTTCAAGAGCA 2382
 837 SerProGluMetSerLeuLysLeuGluLysAsnLeuGluLysProAspAs 853
 2381 TCTCCACAGATGTCAATGAACTCGAAGAGAAATTTAGAGAACCTGATGA 2332
 853 PLeuGluThrGluIleLeuGlnGluProSerGlyThrAsnLysAspGlu 869
 2331 TTTGGAAACAGAAATTTCTACAGAGCCAGGTGAAACAAAGATGAGA 2282
 870 SerLeuProCysThrIleThrAspValTrpIleSerGluGluLysGluTh 886
 2281 AGCTTCCATGCACATTAAGTGTGTGATGTGTGATGTGAGGAAAAAGAAC 2232
 886 rLysGluThrGlnSerAlaAspArgIleThrIleGlnGluAsnGluValS 903
 2231 AAGGAAACTCAAGTCGGCAGATAGATACCATTCAGGAAAAATGAGTTT 2182
 903 etGluAspGlyValSerSerThrValAspGlnLeuSerAspIleHisIle 919
 2181 CTAAATATGAGAGTCTGAGTACTGTGACCAACTAGAGCATTCATATTA 2132
 920 GluProGlyThrAsnAspSerGlnHisSerLysCysAspValAspLysSe 936
 2131 GAGCTGGAGAACATGATTTCTAGACACTGTAAATGTGATGATGATGATG 2082
 936 rValGlnProGluProPhePheHisLysValValHisSerGluHisLeuA 953
 2081 TGTGCAACCGAACCATTTCATTAAGTGTGTGATTCATTCGAAACATTGA 2032
 953 snLeuValProGlnValGlnSerValGlnCysSerProGluGluSerPhe 969
 2031 ACTTAATCCCTCAAGTTCATCAATCAGTTCAGTTCACCGAAGAAATCCTT 1982
 970 AlaPheArgSerHisSerHisLeuProProLysAsnLysAsnLysSe 986
 1981 GCATTTGATCTCAGCTCGCATTTACCAACAAAAATAAACAAGAAATTC 1932
 986 rLeuLeuIleGlyLeuSerThrGlyLeuPheAspAlaAsnAsnProLysM 1003
 1931 CTTGCTGATTTGACCTTCAACTGCTGCTGTTGATGCAAAACCAACCAAGA 1882
 1003 etLeuArgThrCysSerLeuProAspLeuSerLysLeuPheArgThrLeu 1019
 1881 TGTTAAGACATCTTCACCTCCAGATCTCTCAAAAGCTTTTACAGAACCTT 1832
 1020 MetAspValProThrValGlyAspValArgIleAspAsnLeuGluIleAs 1036
 1831 ATGAGATGTTCCCGCTGAGAGATGTTCCTCAAGACAAATCTTGAATAGA 1782
 1036 pGluIleLysAspGluAsnIleLysGluGlyProSerAspSerGluAspI 1053
 1781 TGAATTTGAGATGAAACATTAAAGAGACCTTCTCATTTCTGAACACA 1732
 1053 LeValPheGluGluThrAspThrAspLeuGlnGluLeuGlnAlaSerMet 1069
 1731 TTTGTTTGAAGAACTGACAGAGATTTACAAAGAGCTCAGCGCTCATG 1682
 1070 GluGlnLeuLeuArgGluGlnProGlyGluGluIuTyrSerGluGluGlu 1086
 1681 GAACAGTTACTTAGGAGAACAACTGTGTGAGAAATACAGTGAAGAGAGAGA 1632

1086 uSerValLeuLysAsnSerAspValGluProThrAlaAsnGlyThrAspV 1103
 1631 GTCAGTCTTGAAGAACAGTATGTGAGCCACTGCAGAAATGGAGACAGATG 1582
 1103 alaIAspGluAspAspAsnProSerSerGluSerAlaLeuAsnGluIu 1119
 1581 TGGCAGATGAAGATGACAAATCCAGCAGTGAAGATCCCTGAACCAAGAA 1532
 1120 TrpHisSerAspAsnSerAspGlyGluIleAlaSerGluCysGluCysAs 1136
 1531 TGGCAGTCAAGTACAGTATGTGAAATGCTCAGTGAATGTGAATGCGGA 1482
 1136 pSerValPheAsnHisLeuGluGluLeuArgLeuHisLeuGluGlnGluM 1153
 1481 TAGTGTCTTTACCATTTTACAGAGACTGACACTTCTCGAGAGCAGGAAA 1432
 1153 etGlyPheGluLysPhePheGluValTyrGluLysIleLysAlaIleHis 1169
 1431 TGGCCTTTGAAAAATTTCTTGAAGTTTATGAGAAATTAAGGCTATTCAT 1382
 1170 GluAspGluAspGluAsnIleGluIleCysSerLysIleValGlnAsnI 1186
 1381 GAAGATGACAGTGAATAATTTGAAATTTCTCAAAAATGTTCAAAAATAT 1332
 1186 eLeuGlyAsnGluHisGlnHisLeuTyrAlaLysIleLeuHisLeuValM 1203
 1331 TTTGGAAATGAAACATCAGCATCTTTATGCAAGATTTTCATTTAGTCA 1282
 1203 etAlaAspGlyAlaTyrGlnGluAspAsnAspGlu 1214
 1281 TGGCAGATGAGAGCCTACCAAGAGATATGATGAA 1247
 seq_name: /net/abs06/SIDS/9/cgdata/hold-geneseq/geneseqn-embL/NA2001A.DAT:AA511559
 seq_documentation_block:
 ID AA511559 standard; cDNA; 3024 BP.
 AC AA511559;
 DT 24-OCT-2001 (first entry)
 DE Human cDNA encoding novel human protein, NHP #3.
 XX
 KW Human: novel human protein; NHP; ss; breast cancer; prostate cancer;
 KW Immunogen; antibody; gene therapy; antisense.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3024 /'tag' a
 FT /product= "NHP #3"
 XX
 PN MO200161016-A2.
 PD 23-AUG-2001.
 XX
 PF 15-FEB-2001; 2001WO-US05356.
 XX
 PR 18-FEB-2000; 2000US-0183582.
 PR 22-FEB-2000; 2000US-0184014.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Hu Y, Nepomilchy B, Turner CA, Zambrowicz B;
 XX WPT; 2001-502793/55.
 DR P-PSDB: AA007103.
 XX
 PT Isolated nucleic acids encoding novel human proteins useful for the
 PT treatment of disease and as probes for testing and detection -
 XX
 PS Disclosure: Page 40; 69pp; English.

XX The invention relates to novel human proteins (NHP) and the nucleic acids encoding them. The nucleic acids encode mammalian transporter proteins and are useful for the treatment (e.g. by gene therapy or antisense technology) of any of a wide variety of symptoms associated with biological disorders (e.g. breast and prostate cancer) or imbalances and as probes for the identification, selection and validation of novel molecular targets for drug discovery. The proteins may be used to raise anti-NHP antibodies. The present sequence encodes an NHP of the invention.

Sequence 3024 BP; 1115 A; 503 C; 681 G; 725 T; 0 other:

alignment_scores:
Quality: 5162.00 Length: 1007
Ratio: 5.126 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-783-320-4 x AAS11559

Align seg 1/1 to: AAS11559 from: 1 to: 3024

```
208 MetLysAsnLeuValLeuLysIleIleSerGlySerPheProPvalSe 224
1 ATGAAAAACCTGGTACTGAGATATATCTGGATCTTTTCACCTGTC
224 rLeuHisTyrSerTyrAspLeuArgSerLeuValSerGlnLeuPheLysA 241
51 TTTGGATTATTCCTATGATCTCCGACGTTGGTGCTCACTATTTTAAA
241 rGAsnProArgAspArgProSerValAsnSerIleLeuGlnLysGlyPhe 257
101 GAAATCTCTAGGATGACCATCATCACTCATTTGGAGAAAGGTTT 150
258 IleAlaLysArgIleGlnLysPheLeuSerProGlnLeuIleLagLugl 274
151 ATAGCCAAACGCAATGAAAGTTCTCTCTCAGCTATTCAGAGAG 200
274 upheGlyLeuLysThrPheSerLysPheGlySerGlnProIleProAla 291
201 ATTTTCTTAAACATTTTCGAAGTTGGATCAGCCTATFACACGCTA 250
291 yAsnProAlaSerGlyGlnAsnSerIleSerValMetProAlaGlnLys 307
251 AAGACCCAGCTTCAGACAAACCTGATTTCTGTTATGCCGTCAGAAA 300
308 IleThrLysProAlaAlaLysTyrGlyIleProLeuAlaTyrLysLys 324
301 ATTACAAAGCTGCCCTAAATATGAAATACCTTTAGCATTAAGAAATA 350
324 rGlyAspLysLysLeuHisGlnLysLysProLeuGlnLysHisLysGln 341
351 TGGAGATTAATAAATTACAGAAAGAAACACCTGCANAAACATTAACGG 400
341 IahISglnThrProGlnLysArgValAsnThrGlyGlnLysArgLys 357
401 CCATCAAACTCCAGAGAAAGATGATACGAGAGAGAAAGAGAGAAA 450
358 IleSerGlnLysIleAlaAlaArgLysArgLeuGlnPheIleGlnLys 374
451 ATATCTAGAGACACAGCAAGAAAGAGGCTGGAATTTATTTGAAAGAA 500
374 uLysLysGlnLysAspGlnIleIleSerLeuMetLysAlaGlnLysMet 391
501 AAGAAACAAAGATCAGATTTATTTAGTTAATGAAGCTGACAAATGA 550
391 ySAsnGlnLysGlnArgLeuGlnLysArgIleAsnArgAlaArgGlnLys 407
551 AAGGCAAGAAAGAAAGAGTTGGAAGATTAATAGGCCGAGGAAACAA 600
408 GlyTrpArgAsnValLeuSerAlaGlyGlySerGlyValLysAlaPr 424
```

```
601 GGATGAGAAATGTGCTAAGCTGCTGAGAGGAGGAGTAAAGCTCC 650
424 PheLeuGlnSerGlyGlyThrIleAlaProSerSerPheSerArgG 441
651 TTTTCTGGGCGAGTGGAGGAGCTATACCTCCATCATCTTTTCTTCAG 700
441 yGlnTyrGlnLysIleThrHisAlaIlePheAspGlnMetGlnGlnArg 457
701 GACAGTATGAACATTTACCATTCCTTTTTCACCAATTCAGCAACAGAA 750
458 AlaGlnAspAsnGlnAlaLysTyrPysArgGlnIleTyrGlyArgGly 474
751 GCAGAGATTAATGAAGCTTAATGAGAAAGAAATATATGCTGAGGCT 800
474 uProGlnArgGlnLysGlnLeuAlaValGlnLysGlnLysGlnValG 491
801 TCCAGAAAGGCAAAAAGGCGAGCTGCTGAGAAAGAGCTAAACAGAG 850
491 LuGlnPheLeuGlnArgLysArgGlnLysMetGlnAsnLysAlaArgAla 507
851 AAGAGTCTCTGAGCGAAACGGAGACCTATGAGATTAAGCTGAGGC 900
508 GlnGlyHisMetValTyrLeuAlaArgLeuArgGlnIleArgLeuGln 524
901 GAAGGACATATGCTTATCTGCAAGACTGAGCAATAAAGACTACAGAA 950
524 nPheAsnGlnArgGlnGlnIleLysAlaLysLeuArgGlyGlyLysG 541
951 TTTCAATGAGCCCAACAGATTAAACCAAACTCGGTGGAAGAAAG 1000
541 LuAlaAsnHisSerGlnGlnGlnGlnLysSerGlnGlnLysAlaPhe 557
1001 AAGCTATATCTCTGACAGACAGAGAGAAAGTGAAGGCTGACATGAG 1050
558 ArgLysLysIleGlnSerLeuLysAlaHisAlaAsnAlaArgAlaLys 574
1051 CGCAAAAAATGCAATCATGAGGCCCATCAATGACAGTCTCTCTG 1100
574 IleLysGlnGlnLeuGlnLysArgLysGlnLysAlaTyrGlnLysGln 591
1101 ACTTAAAGAAACAACAGAGAGAAAGAGAGAGCTTATGAGAGAGAAA 1150
591 yLysValTyrPglGlnLysIleLeuValAlaLysGlyValLysSerSer 607
1151 AAAAAGTGTGGAGAGCAATTTGTGGCTAAAGAGTTAAGATTCGAT 1200
608 ValSerProProLeuGlnLysIleSerValThrSerAlaLeuLysGln 624
1201 GTTCTTCCACCTTTGGACAGCATGAACAGGTGGCTCTCCATCAAGCA 1250
624 nGlnMetArgSerValIleSerValThrSerAlaLeuLysGlnLys 641
1251 ACAGATGAGATCTGTAATTTCTGTAACCTTCAGCTTGAAGAAAGTGG 1300
641 aLAspSerSerLeuThrAspThrArgLysThrSerGlnLysMetGlnLys 657
1301 TGGACACTAGTTTAACGTATACCGGAAACTTCAGAGAGATCAAAAG 1350
658 ThrAsnAsnAlaIleSerSerLysArgGlnLysLeuAlaArgLysGln 674
1351 ACCAACAATGCTATTTCAAGTAAGAGAAATACCTTCACATTAATGA 1400
674 uAsnLeuLysAlaGlnLysAspGlnLysGlyMetGlnAsnLeuSerSPT 691
1401 AAATCTTAAGCTCAAGAGATGAAGAAAGAAATGCAAGATCTCTGTATA 1450
691 hrPheGlnIleAsnValHisGlnAspAlaLysGlnHisGlnLysGlnLys 707
1451 CTTTGGATTAATGTTCAATGAGATGCAAGAGAGATGAAGAAAGAA 1500
708 SerValSerSerAspArgLysLysTyrPglLysAlaGlyGlnLeuVal 724
```


341	LaHsgnThrProGluLysAcroValAspThmGlyGlnGluArgArgLys	357
401	CCCTCAACTCCAGAGAAAGAGCTGATATCTGCAAGAAAGAGGAAA	450
358	IleSerGluGluAlaAlaArgLysArgArgGluGluPheIleGluLysG	374
451	ATATCTGAGGAAGCGCAAGAAAGAGAGGCTGGAAATTATTCAGAAAAGA	500
374	uLysLysGluLysAspGlnIleIleSerLeuMetLysAlaGluGlnMetL	391
501	AAACAAACAAAGAGATCAGTTATTATGTTTAAAGAAGCGTGAAACAATA	550
391	LysArgGlnGluLysGluArgLengLuuArgLleAsnArgAlaArgGluGln	407
551	AAAGCAAGAAAGAGAAAGCTGTGAAAGAAATTAATGGCCAGAGGAAACA	600
408	GlyTPArgAsnValIleuSerAlaGlyGlySerGluValLysAlaBr	424
601	GGATGTGAAGAAATGTGCTTAAGGCTGGTGGCAAGCTGTAAGTAAAGGCTCC	650
424	OphLeuGluLysSerGlyGlyThrIleAlaProSerSerPheSerSerArg	441
651	TTTTTGGGCGAGTGGAGGAGCATAGCTCCATCATATCTTTTCTTCGAG	700
441	LysGlnTyrGluHisTyrHisAlaIlePheAspGlnMetGlnGlnGlnArg	457
701	GACACTATGACACATTATCCATGCCATTTTTGACCAATATGCCAGCAACAAAGA	750
458	AlaGluAspAsnGluAlaLysTyrLysArgGluLysTyrGlyArgGlyLe	474
		800

474 uProGluArgGlnLysGlycInLeuAlaValGluArgAlaLysGlnValG 491
 |||||
 801 TCCGAAAGGCCAAAAAGGCCAGCTAGCTAGAAAGGCTTAACCACTAG 850

[illegible]

508 GlucylhistMet..... 511

901	GAAGGCATATGGGAATCTGCCAAACCTGGCAGCTATGTATGGAGCAG	950
512	ValT 513
951	GCCCACTCTTCAAGAGGAGGCAACCAAGAACAAAGAGAGAGGTTT	1000
513	yrleuAlaArgleuArvglnileatcgleuGlnPheasnGluArgln	529
1001	ATCTGGCAAGACTGAGGCCAAATTAAGACTCAACAATTCATAGAGGGCCAA	1050
530	GlnIleIysAlaIysleuArgIyGlnIuIyIysGlnAlaasnHisSerGI	546
1051	CAGATTAAACCCAAACCTCTGTGTGAAAAGAAAGAGCTAATCATTCTGCA	1100
546	uGIyGlnGluGIySerGIuGluAlaaspMetArgIyIysIleGluS	563
1101	AGGCACAAGAGGAAGTGAAGAGGCTGCATGTAGGGCCAAAAAATCGAAT	1150
563	erleuIysAlaHisAlaasnAlaArgAlaAlaValleuIysGluInleu	579
1151	CACGTGAAGGCCCATGCAATGCACGTCGTCTACTAAACAACAACATA	1200
580	GluArgIyAspArgIyGlnAlaIyArgIyGlnAlaIyValIleTrpGlnI	596
1201	GAACGAAAGGAAAGGAGGCGCTTATGAGAGGAAAGAAAAAGTGTGGAGAA	1250
596	uHisIeuValAlaIySGIyValIySerSerAspValSerProIoleuG	613
1251	GCATTTGTGTGCTTAAGAGATTTCTGATGTTCCTCCACCTTTGG	1300
613	IyGlnHisGluThrGIyGIySerProSerIySGlnInleArgSerVal	629

1301	GAAGCATGAAACAGCTGCTCTCCATTCACAAAGCAACAGATGACATCTT	1350
1302	GAAGCATGAAACAGCTGCTCTCCATTCACAAAGCAACAGATGACATCTT	1350
630	IIeserValThrSerAlaLeuysGluValGlyValAspSerSerLeuTh	646
1351	ATTTCGTACCTTCAGCTTTTCAAGAAAGTTGGCTGGACAGTAACTTAC	1400
646	rsaphrThrArgLutThrSerGluGluMetGlnlyStrAsnAlaIles	663
1401	TGATACCCGGGAAACTTCAGAAAGATGCAAAAGACCAACAATGCTATT	1450
663	erSerLysArgGluIleLeuArgArgLeuAsnGluAsnLeuysLacIn	679
1451	CAAGTAAAGGAGAAATACTTCGGCAGATTAAATGAATAATCTTAAACCTAA	1500
680	GluAspGluLysGlyMetGlnAsnLeuSerAspThrPheGluIleAsnVa	696
1501	GAAGATGAAAAAGAAATCCAAATCTCTCGATATCTTTGAGATAAAGT	1550
696	IIAsGluAspAlaLysGluIleGluLysGluLysSerValSerSerAspA	713
1551	TCATGAAAGTCCCAAGAGCATGAAAAAGAAATCAGATTTCATCTGATC	1600
713	rgLysLysTrpGluLacGlyGlyGlnLeuValIleProLeuAspGluLeu	729
1601	GCAGAGAGGGGAGAGCAGAGCTCACTTGATCTCTGGATGAGTTA	1650
730	ThrLeuAspThrSerPheSerThrThrGluArgHisThrValGlyGluVa	746
1651	ACACTGATACACTCTCTCTACAACTGAAAGACATACAGTGGAGAGT	1700
746	IIleLysLeuGlyProAsnGlySerProArgArgAlaTrpGlyLysSerP	763
1701	TATTAAATAGGTCTTAATGATGATCTCCAAAGACCTGGGGAAAAATC	1750
763	roThrAspSerValLeuLysIleLeuGlyGluValGluLeuInleuIn	779
1751	CGACAGATTCTGTTTAAAGACTCTGGAGAGCTACAACTACAACTTCA	1800
780	ThrGluLeuLeuGluAsnThrThrIleArgSerGluIleSerProGluG	796
1801	ACGAGACTTTAGAAATATACAACTATTAGAGAGATGATTTCCCGGAAG	1850
796	yeLutLysTrpLysProLeuIleThrGlyGlyLysLysValGlnCysIles	813
1851	GGAAAAAGTACAAACCTTAAATTAATCTGGAGAAAAAAAGTACAAATATT	1900
813	erHisGluIleAsnProSerAlaIleValAspSerProValGluThrLys	829
1901	CACATGAAATAAACCATACACTATCTGTCATCTCTTACAGCAAAA	1950
830	SerProGluLysSerGluLysLysProGlnMetSerLeuLysLeuGluG	846
1951	AGTCCCGAATTCAGTGAAGCATCTCCACAGATGTCAATGAACCTGGAAG	2000
846	ysnLeuGluGluProAspAspLeuGluThrGluIleLeuGlnGluProS	863
2001	AAATTAGAGAACCTGATGATTGGAAACAGAAATTTCTACAAAGCCAA	2050
863	erGlyThrAsnLysAspGluSerLeuProCysThrIleThrAspAlaTrp	879
2051	GTCGAAACAAACAAAGATGAGCTTGGCATGCGCACTTACTGATGTGG	2100
880	IleSerGluGluLysGluThrLysGluThrGlnSerAlaAspArgIleTh	896
2101	ATTAGTGAAGAAACAAACAAAGAACTCACTGGCAGATGAGATAC	2150
896	rlIleGlnGluAsnGluValSerGluAspGlyValSerSerThrValAsp	913
2151	CATTCAGGAAAAATGAAGTTCTTGAAAGTGAAGTCTGAGTACTGTGAGC	2200
913	IIleuSerAspIleHisIleGluProGlyThrAsnAspSerGlnHisSer	929
2201	AACTTAGTCACATTCATTAAGCTCTGCAACCAATGATTTCTGACAGCT	2250
930	LysCysAspValAspLysSerValGlnProGluProPhePheHisLysVa	946
2251	AAATGATGATGATTAATCTGTCTGCAACCGCAACATTTTTCATTAAGT	2300
946	IValHisSerGluHisLeuAsnLeuValProGlnValGlnSerValGlnC	963
2301	GGTTCATTTCTGAACCTTGAACCTTAAGTCCCTCAAGTTCAATCAGTTCA	2350
963	ysSerProGluGluSerPheAlaPheArgSerHisSerHisLeuProPro	979
2351	GTTCCACGAGAAAGAAATCTTTCATTTGCAATCTACATCTGCAATTTACCA	2400
980	LysAsnLysAsnLysAsnSerLeuLeuIleGlyLeuSerThrGlyLeuPh	996
2401	AAAAATAAAAACAAATATCTTCTGATGAGCTTTCACAACTGCTCTGT	2450
996	eaSPAlaAsnAsnProLysMetLeuArgTrpCysSerLeuProAspLeus	1013
2451	TGATGCAAAACCAACCAAGATGTTAAAGAAATGTTCACTTCCAGATCT	2500
1013	erLysLeuPheArgThrLysMetLysValProThrValGlyLysValArg	1029
2501	CAAACTGTTTCAGAACCTTATGATGTTTCCACCCTAGAAATGTTTGT	2550
1030	GlnAspAsnLeuGluIleAspGluIleLysAspGluAsnIleLysGluG	1046
2551	CAAGACAACTCTGAAATAGATGAAATTAAGATGAAACATTTAAAGAG	2600
1046	yProSerAspSerGluAspIleValPheGluGluThrAspThrAspLeu	1063
2601	ACCTTGTGATTTGAAAGCATTTGTTTGAAGAACTGACACAGTTTAC	2650
1063	IngLysLeuGlnAlaSerMetGluLeuLeuLeuArgGluLysProGluG	1079
2651	AAAGCTGAGAGCTCGATGGAACAGTTACTTTGGGAACAACTGCTGTA	2700
1080	GluTrpSerGluGluGluGluSerValLeuLysAsnSerAspValGluP	1096
2701	GAATACAGAGAAAGAAAGATCACTTCTTGAAGAACAGATGATGAGCC	2750
1096	oThrAlaAsnGlyThrAspValAlaAspGluAspAspAsnProSerSerg	1113
2751	AACTGCAAATGGGACAGATGTGGCAGTCGAAGTGAACAATCCCACTA	2800
1113	IuSerAlaLeuAsnGluLutTrpHisSerAspSerSerpLysGluIle	1129
2801	AAAGTCCCTCGAACCAAAATAGGCACACTCAAGATACAGTATGCTGTA	2850
1130	AlaSerGluCysGluCysAspSerValPheAsnHisLeuGluLysLeuAr	1146
2851	GCTAGAGATGTGAAATGCGATAGCTGTTTAAACATTTAGAGAACTGAG	2900
1146	gLeuHisLeuGluGlnGluMetGlyPheGlyLysPhePheGluValTyg	1163
2901	ACTTCATCTCGAGAGAAATGGGCTTTGAAAAATCTTTGAGGTTTATG	2950
1163	IuLysIleLysAlaIleHisGluAspLysAspLysGlnIleGluIleCys	1179
2951	AGAAAAATTAAGCTATTCTTAATAATAATAATAATAATAATAATTTCT	3000
3001	TCAAAAATATGTTTATTTTGGAAATGAACATACACAGATCTTTATAGC	3050
1180	SerLysIleValGlnAsnIleLeuGlyAsnGlnHisGlnHisLeuVal	1196
3051	CAAGATTTCTCATTTAGTCATGCGCAGATGAGCCTTACCAAGAAAGTAA	3100
1196	alysIleLeuHisLeuValMetAlaAspGlyAlaTygGlnLysAspAsn	1213
1213	spGlu 1214	
3101	ATGAA 3105	

seq_name: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AAA09328

```
seq_documentation_block:
ID      AAA09328 standard; DNA; 4263 BP.
```

AC AAA09328;

DT 10-AUG-2000 (first entry)

DE Human cancer associated antigen precursor DNA, clone NY-REN-55.

KM renal cancer; cancer associated antigen precursor; diagnosis
 KM cytostatic; murine NEK1 protein kinase homologue; ss.

OS Homo sapiens.

PN WO200020587-A2

PD 13-APR-2000

PF 04--OCT-1999; 99WO-US22873

PR 05-OCT-1998; 98US-0166300.

[illegible]

AA
PA
(LUDW-) LUDWIG INST CANCER RES.

Obata Y, Gout I, Tureci O, Sahin U, Pfreundschuh M, Scanlan MJ;

PI Stockert E, Chen Y, Old LJ, Jager E, Knuth A

WPI; 2000-303774/26.

aa Preventing, diagnosing and/or treating disorders associated with
 pt abnormal expression of human cancer associated antigens

aa
PS Claim 57; Page 100-101; 121pp; English

AA09321.43 were isolated by SEEX screening from a renal cancer cell line 1773/10.4. Homology searching revealed that these clones correspond to known genes. The present sequence has identity with the murine NFK1 protein kinase homologous gene. The genes encode cancer associated antigen precursors. These genes are useful in methods for preventing, diagnosing and/or treating disorders, especially cancers. The method comprises contacting a sample from a subject with an agent that specifically binds to the nucleic acid molecule of expression product (or fragment) complexed with a human leukocyte antigen (HLA) molecule and determining the interaction between the agent and the nucleic acid molecule or the expression product as a determination of the disorder.

Sequence 4263 BP; 1364 A; 829 C; 1065 G; 1005 T; 0 other:

alignment_scores:

Quality:	5040.50	Length:	1214
Ratio:	4.496	Gaps:	6
Percent Similarity:	92.339	Percent Identity:	81.796

alignment_block:

Align seg 1/1 to: AAA09328 from: 1 to: 4263

1 MetGlnLysrYrAlaArfLeuGlnLysrIleGlyIugIuLysrPheGly 17
576 ATGAGACGATGTGTGACATGCACGACGATTTGGACAAAGTTCATTGGAAA 622

[illegible]

1576 CACCCCAAAACATMAACAGCCCATCAATTCCTCGAAGAAAGAAAT 1625
1576 ThrGluGluGluArgArgLysIleSerGluGluAlaAlaArgLysArg 367
1576 TCTGAGAGAGAAAGAGAAATGCTGAGAGACGACCAAAAAAGAG 1675
1576 gLengluPheIleGluLysGluLysGluLysAspGluIleIleSerL 384
1576 GTTGAATTTATGTAGAAAGAAAGCAAAAGATCAG...ATTAGGT 1722
384 eumElysAlaGluGluMetLysArgGluGluLysGluLysArgLys 400
1723 TCTGAAGGCTGACAGATGACGCGACAGAGACGCGTTGGAGAG 1772
401 IleAsnArgAlaArgGluGluGlyTPArgAsnValLeuSerAlaGly 417
1773 ATTAATAGGCGCAGGAAACAGAGATGGAAATGTTTAAAGGCTGTGG 1822
417 ySerGlyGluValLysAlaProPheLeuGlySerGlyGlyThrIleAla 434
1823 AAGCGTGAATTAAGGCTTCCTTTTGGCATTTGAGAGGCGCTGCTCTC 1872
434 roSerSerPheSerArgGlyGluGlyGlyIleIleIleIleIleIle 450
1873 CATCACCTGTCTCTCTCGAGCCAGATGACATTACATTCATTCATTTT 1922
451 AspGluMetGluGluGluArgAlaGluAspAsnGluAlaLysTrpLys 467
1923 GACCAATGCGACGCGTAAAGAGCAAGATTAAGCAAGCAAGAGAGG 1972
467 gGluIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 484
1973 GGGATCTATGCTGATGCTGCCAGAAAGCAAAAGGCACTTATGCTG 2022
484 aGluArgAlaLysGluValGluGluPheLeuGluGluArgLysArg 500
2023 TAGAGAGAGCCACCAAGTGAAGAAATCTTACAGCTAAACAGAGAG 2072
501 MetGluAsnLysAlaArgAlaGluGlyHisMetValTrpLeuAlaArg 517
2073 ATGCGAATTAAGCCGAGCGAGAGACAGCTGTTATTTGGCAAGACT 2122
517 uArgGluIleArgLeuGluAsnPheAsnGluArgGluGluIleLysAla 534
2123 GAGCAATTAAGACTACAAAATTTTAATGAGCCCAACAGATTAAGCCA 2172
534 ySLeArgGlyGluLysLysGluAlaAsnHisSerGluGluGluGly 550
2173 AACTTCGCTGAGATTAAGAGAGCTGATGACCAAGAGACAAAGAGCA 2222
551 SerGluGluAlaAspMetArgArgLysLysIleGluSerLeuLysAla 567
2223 ACTGAAGAGACTGACATGAGCTCAAAAAGATGAGTCACTTAAGCGCA 2272
567 sAlaAsnArgAlaAlaValLeuLysGluGluGluGluGluGluGlu 584
2273 AACAAATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2322
584 ySGluAlaTrpGluArgLysLysValTrpLeuGluHisLeuValAla 600
2323 AGGAGGCTTATGAAGAGAAAGAAATGATGGAAGAACTTTGGTGCGG 2372
601 LysGluValLysSerSerAspValSerProLeuGluGluHisGluThr 617
2373 AGG...GTAAGAACTCAGATGTTCTCTCTGCTTGGAACTTCTTAAAC 2419
617 rGlyGlySerProSerLysGluGluMetLysValIleSerValHis 634
2420 AGGTGCTTCTCATCAAGAGAGAGAGCTGCTCATTTCTGCTGCTGCT 2469
634 eAlaLeuLysGluValGluValAspSerSerLeuThrAspTrpArgGlu 650

2470 CAGCTTTGAAGAGAGTGGCTGATGGAAGTTTAAGTATACCAAGAA 2519
651 ThrSerGluGluMetGluLysThrAsnAlaIleSerSerLysArgL 667
2520GAAGAAATGAGAAAGAGTAACAGTCTATTTCAAGTAAAGCA 2563
667 uIleLeuArgArgLeuAsnGluAsnLeuLysAlaGluGluLysGlu 684
2564 AATCTGCGGAGCTTAATGAATCTTAACCTCAAGAGATGAGAAAG 2613
684 LysMetGluAsnLeuSerAspThrPheGluIleAsnValHisGluAsp 700
2614 AAAGAGCATCTCAGTCTTGTGAGACCGCTTGCTGACAAAGATGAG 2663
701 LysGluHisGluLysGluSerValSerSerAspArgLysTrpGlu 717
2664 AGAGAGATGAGACAGAAATGCTATTCCTGATGCTCAAGAGATGGA 2713
717 uAlaGlyGluGluValIleProLeuAspGluLeuThrLeuAspThr 734
2714 GATGGAGGCTCAGCTGTGATGCTCTGATGCTGATGCTGATGAT 2763
734 ePheSerThrThrGluArgHisThrValGlyGluValIleLysLeu 750
2764 CTTTCTCTGACCCGAAACATACCTGAGAGAGTATTAATTAATGAT 2813
751 ProAsnGlySerProArgArgAlaTrpGlyLysSerProThrAspSer 767
2814 TCTAATGCTCTCCAGAAAGCTGGGGAAGAAACCTTACAGATTCGT 2863
767 IleLysIleLeuGluGluAlaGluLeuGluGluGluGluGluGlu 784
2864 GCTGAAGATACCTGGAAGAGTGAATTAACCTATGACAGACAGACTAG 2913
784 LysAsnThrThrIleArgSerGluLysSerProGluGluLysTrpLys 800
2914 AAACACATCTTTAAAGTGAAGTTATGCTGAAGAGAGACATCTCA 2963
801 ProLeuIleThrGluLysLysValGlnCysIleSerHisGluIleAs 817
2964 CCCTTCTTACAGAGAGAGATGCTGAGTCAATTTCAAAAGAAATTA 3013
817 nProSerAlaIleValAspSerProValGluThrLysSerProGluPhe 834
3014 TCCATCAGCTACCTGTTGATTTCT...ACTGAAGCAAAAGTCCAAAGTTA 3060
834 eGluAlaSerProGluMetSerLeuLysGluGluLysLeuGluGlu 850
3061 CTGAGGCTGCTCCCAAAATGTCA.....GAAGGAAATGTGAGAGAA 3101
851 ProAspAspLeuGluThrGluIleLeuGluGluProSerGlyThrAsn 867
3102 CTTGATGATTTGGAACAGAGATTTACAAAGAGCCAAATAGACACAC 3151
867 sAspGluSerLeuProCysThrIleThrAspValTrpIleSerGluGlu 884
3152 AGATGGAGATTTCCACCTGTTCTTAATGATGTGCTGAGTGAAGAG 3201
884 ySGluThrLysGluThrGluSerAlaAspArgIleThrIleGluGlu 900
3202 AACGAGCTAAGAACTGAGTTGGAAGATAAAGTCTGCTGAGAGAGT 3251
901 GluValSerGluAspGlyValSerThrValAspGluLeuSerAsp 917
3252 GAAGTTTGCAAGATGAATTCACAGGAACGTGACCAATCCTGAAGCA 3301
917 eHisIleGluProGlyThrAsnAspSerGluHisSerLysCysAsp 934
3302 TCAGAGAGATCTCTGAGTACAGATTCCTGCGACGCTGCTGATGAG 3351
934 sPlySerValGlnProGluProPhePheHisLysValValHisSerGlu 950
3352 AGAAGTCAGTACAGCCAGAAATGATTTTCCAAAGAGTGTTCATTCAG 3401


```

951 HisLeuAsnLeuValProGlnValGlnSerValGlnCysSerProGlnGlu 967
    |||||
3402 GACTTGAACCTTA.....GTTACGACGAGCTTCATCTCCACACGAGA 3442
967 userPheAlaPheArgSerHisSerHisLeuProGlnAsnLysAsnL 984
    |||||
3443 ACCAATTCCAAATTCGATCTCACTGATTCCTCCACCAAACTAAGAGCA 3492
984 ysaAsnSerLeuLeuIleGlyLeuSerThrGlyLeuPheAspAlaAsn 1000
    |||||
3493 AGAATTCCTTACGATTCGACCTTCAACTGCTGCTTTGATGACAAACAT 3542
1001 ProlysettleuArgThrCysSerLeuProAspLeuSerLysLeuPheAr 1017
    |||||
3543 CCAAAAGTCTGAGGAGACCTGCTCAGATTCCTTCCAAAGCTGTGAC 3592
1017 gThrLeuMetAspValProThrValGlyAspValArgGlnAspAsnLeuG 1034
    |||||
3593 AACCTTATGAGACCTTCCACTGTGGGACCTTCATCAAGACAGCTCTTG 3642
1034 IuileAspGluIleLysAspGluAsnIleLysGlnGlyProSerAspSer 1050
    |||||
3643 AAATCGATGAGCTGGAAGATGAAACCAATTAAGAGGGCTTCTGATTC 3692
1051 GluAspIleValPheGlnGluThrAspThrAspLeuGlnLeuGlnAl 1067
    |||||
3693 GAGACACCTGATATTGAAAGAACTGACACAGATTTCACAGACCTTCAG 3742
1067 aSerMetGlnGlnLeuLeuArgGlnGlnProGlyGlnGlySerGlnG 1084
    |||||
3743 CCAATGAGACGACCTGCTAGGGGACCAACAGGTGAGCAATACAGTGA 3792
1084 IuGlnGluSerValLeuLysAsnSerAspValGluProThrAlaAsnGly 1100
    |||||
3793 AGCAGAGACTGCTGTTTAAAGACGAGCTGTGAGCAGACACACAGAGG 3842
1101 ThrAspValAlaAspGluAspAspAspProSerSerGlnSerAlaLeuAs 1117
    |||||
3843 ACAGATGCGCCACGACGAGAGACACCCACGACGCAAAAGCC.CTGAA 3891
1117 nGlnGluThrPheHisSerAspAsnSerAspGlyGluIleAlaSerGluCysG 1134
    |||||
3892 CGAGGATGGGACCTGACATATAGTACGCGTACAGACCACTAGATGTG 3941
1134 IuGlyAspSerValPheAsnHisLeuGlnGluLeuArgLeuHisLeuGln 1150
    |||||
3942 AATATGACAGTCTCTTAAACATTTAGAGAACTAAAGACTTCACTGGAG 3991
1151 GlnGluMetGlyPheGlnLysPhePheGlnValArgLysIleLysAl 1167
    |||||
3992 CAGAAATGGCTTGAAGAGCTCTTGAGGTTTATGAGAAAGTAAAGGC 4041
1167 aIleHisGluAspGluAspGluAsnIleGluIleCysSerLysIleValG 1184
    |||||
4042 TATCTATGAGGATGAGATGAAATATGAAATTTGTTCAACATAGTTG 4091
1184 InAsnIleLeuGlnLysGlnLysHisLeuThrAlaLysIleLeuHis 1200
    |||||
4092 AGAATATTTTGGCAATGAGCACGACATCTCTATGCCAACAAATCTGCAT 4141
1201 LeuValMetAlaAspGlyAlaLysArgGlnLysAspAsnGlu 1214
    |||||
4142 TTAGTCATGCGACATGAGCCTATACAGAGATATATGATGAA 4183

```

```

seq_name: /net/abs06/SIDS1/gcgdata/hold-sequences/geneseg-emb1/NA2001A.DAT:AAH17731
seq_documentation_block:
ID AAH17731 standard; cDNA: 2975 BP.
XX AAH17731:
XX
XX 26-JUN-2001 (first entry)

```

```

XX Human cDNA sequence SEQ ID NO:17341.
DE
XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
KW
XX Homo sapiens.
XX BP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000: 2000EP-0116126.
XX
XX 29-JUL-1999: 99JP-0248036.
XX
XX 27-AUG-1999: 99JP-0300252.
XX
XX 11-JAN-2000: 2000JP-0118776.
XX
XX 02-MAY-2000: 2000JP-0183767.
XX
XX 09-JUN-2000: 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 17341; 2537PP + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH13632
XX AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX Sequence 2975 BP; 894 A; 556 C; 562 G; 963 T; 0 other;

```

```

alignment_scores:
Quality: 2012.00 Length: 387
Ratio: 5.199 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-783-320-4 x AAH17731
Align seq 1/1 to: AAH17731 from: 1 to: 2975
91 ArgIleAsnAlaGlnLysGlyValLeuPheGlnGluAspGlnIleLeuAs 107
|||||
3 CGAATTAATGCTCGAAGAGCGTTTGTTCAGAGGATCAGATTGGA 52

```

```

107  pTTPheValGlnIleCysLeuAlaLeuLysHisValHisAspArgLysI 124
108  |||||
109  53  CTGGTTTACGAGATATGTTGGCCGTGAACATGATGATGATGAAAA 102
110  |||||
111  124  LeuHisArgAspIleLysSerGlnAsnIlePheLeuThrLysAspGly 140
112  |||||
113  103  TTTCTTCATCGAGACATTAATCTCAGAAACATATTTTAACATAAGATGGA 152
114  |||||
115  141  ThrValGlnLeuGlyAspPheGlyIleAlaArgValLeuAsnSerThrVal 157
116  |||||
117  153  ACAGTACAAACTGGAGATTTTGGAAATGCTAGACTCTTAATAGTACTGT 202
118  |||||
119  157  GGUleuAlaArgThrCysIleGlyThrProTyrTyrLeuSerProGlnI 174
120  |||||
121  203  AGAGCTGGCTCGAACTTCATAGGGACCCCACTACTCTTCACCTGAAA 252
122  |||||
123  174  LecGlyAsnLysProTyrAsnAsnLysSerAspIleTTPAlaLeuGly 190
124  |||||
125  253  TGTGTAACAAACACCTTACATATATAAGTGACATTTGGCTGTGGG 302
126  |||||
127  191  CysValLeuTyrGlyLeuCysThrLeuLysHisAlaPheGluAlaGly 207
128  |||||
129  303  TGTGTCTTTATGAGCTGTGTACACTTAACATGCTTTTGAACCTGGCAG 352
130  |||||
131  207  fMetLysAsnLeuValLeuLysIleIleSerGlySerPheProProVal 224
132  |||||
133  353  TATGAAAAACCTGTACTGAGATATATATCTGGATCTTTCCACCTGTGT 402
134  |||||
135  224  eTLeuHisTyrSerTyrAspLeuArgSerLeuValSerGlnLeuPheLys 240
136  |||||
137  403  CTTCATTAATTCATATGCTCCGAGTTGGTGCTCAGTTATTTAA 452
138  |||||
139  241  ArgAsnProArgAspArgProSerValAsnSerIleLeuLysGlyLys 257
140  |||||
141  453  AGAATCTCTAGCGATAGACCACTCACTCACTCCATTTGGAGAAAGTTT 502
142  |||||
143  257  eTLeuAlaLysArgIleGlyLysPheLeuSerProGlnLeuIleAlaGly 274
144  |||||
145  503  TATACCCAAAGCGATTGAAAGTTTCTCTCTCCCTCACTTATTTGCGAAG 552
146  |||||
147  274  fMetLysLeuLysThrPheSerLysPheGlySerGlnProIleProAla 290
148  |||||
149  553  AATTTTGTCTAAAAACATTTTCGAAGTTTGATCAGCTATATACAGCT 602
150  |||||
151  291  LysArgProAlaSerGlyGlnAsnSerIleSerValMetProAlaGln 307
152  |||||
153  603  AAAAGACCCAGCTTCAGACAAAACTCGATTTCTGTATCGCTCGTCAAG 652
154  |||||
155  307  sTLeuThrLysProAlaAlaLysTyrGlyIleProLeuAlaTyrLysLys 324
156  |||||
157  653  AATTACAAAGCCTGCCCTAAATATGGAATACCTTATGCAATATATAGAA 702
158  |||||
159  324  yTGLysPheLysLysLeuHisGlyLysLysProLeuGlnLysHisLysGln 340
160  |||||
161  703  ATGGAGATTAATAAATTACACGAAAGAAACACCTGCAAAAACATTAACAG 752
162  |||||
163  341  AlaHisGlnThrProGlyLysArgValAsnThrGlyGlyGluLysArgGly 357
164  |||||
165  753  GCCCATCAAACTCCAGAGAGAGAGTGAATCTGGAGAAAGAGAGGAA 802
166  |||||
167  357  sTLeuSerGlnGluAlaAlaArgLysArgArgLeuGluPheIleGluLysG 374
168  |||||
169  803  AATATCTAGAGAAAGCAGCAAGAAAGAAAGAGCTGGAAATTTATGAAAGG 852
170  |||||
171  374  fMetLysLeuGlnLysAspGlnIleIleSerLeuMetLysAlaGluGlnMet 390
172  |||||
173  853  AAAAAGAAACAAAGAGATCAGATTTATGTTTAAGAGAGCTGAAACAATG 902
174  |||||
175  391  LysArgGlnGlyLysGluArgLeuGluArgIleHisAspArgAlaArgGlu 407
176  |||||
177  903  AAAAGCGAAAGAAAGAGAGCTTGGAAAGATTAATAGGCGCAAGGAGAAC 952
178  |||||
179  407  nGlyTTPArgAsnValLeuSerAlaGlyGlySerGlyGluValLysAlaP 424

```

```

|||||
953  AGCATGGAGAAATGCTTAAGTCTGAGTGAAGTGTGAAGTAAAGGCTC 1002
424  roPheLeuGlySerGlyGlyThrIleAlaProSerSerPheSerSerArg 440
1003  CTTTTCTGGCCAGCTGAGGAGACTATGCTCCATCATCTTTTCTCTCGA 1052
441  GUGlnTyrGlnHisTyrHisAlaIlePheAspGlnMetGlnGlnAla 457
1053  GCACAGTATGACATTACCATTCATTTTGGACCAAAATGCACACAAAG 1102
457  gAlaGluAspAsnGluAlaLysTyrLysArgGlyIleTyrGlyArgGly 474
1103  AGCGAAGATATATCAACCTTAATGAGAAAGAAATATATATGTCGAGCTC 1152
474  euProGluArg 477
1153  TTCGAGAAAGA 1163

seq_name: /net/abs06/SIDSL/ycgdata/hold-geneseg/geneseg-emb1/NA2000.DAT:AAZ86797
seq_documentation_block:
ID   AAZ86797 standard; DNA; 1846 BP.
XX
AC   AAZ86797:
XX
DT   17-APR-2000 (first entry)
XX
DE   Human protein kinase homologue coding sequence, PKH-6.
XX
KW   protein kinase homologue; human; PKH; diagnosis; therapy; cancer; AIDS;
KW   autoimmune disorder; inflammatory disorder; reproductive defect; asthma;
KW   diabetes mellitus; infertility; ovulatory defect; endometriosis;
KW   polycystic ovary syndrome; ss.
XX
OS   Homo sapiens.
XX
PN   US6013455-A.
XX
PD   11-JAN-2000.
XX
PF   15-OCT-1998; 98US-0173581.
XX
PR   15-OCT-1998; 98US-0173581.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Hillman JL, Yue H, Yang YT, Corley NC, Gorgone GA, Azimzal Y;
PI   Lu DAM, Bandman O, Guegler KJ;
DR   WPI: 2000-136321/12.
DR   P-PSDB: AAT76753.
XX
PT   Nucleic acids encoding a human protein kinase homolog useful for
PT   preventing, diagnosing and treating cancer, autoimmune/inflammatory
PT   disorders and reproductive defects -
XX
PS   Claim 6: Column 65-68; 38pp; English.
XX
CC   This sequence encodes a human protein kinase homolog (PKH) of the
CC   invention. The PKH sequences may be used in the prevention, treatment and
CC   diagnosis of diseases associated with inappropriate PKH expression such
CC   as cancers, autoimmune/inflammatory disorders and reproductive defects.
CC   They may be used to treat disorders associated with decreased PKH
CC   expression such as cancers (e.g. lymphoma, melanoma and cancers of the
CC   breast lung and prostate), autoimmune/inflammatory disorders
CC   (e.g. AIDS, asthma and diabetes mellitus), and reproductive
CC   defects (e.g. infertility, ovulatory defects, endometriosis and
CC   polycystic ovary syndrome). The DNA may be administered to treat diseases
CC   by rectifying mutations or deletions in a patient's genome that affect
CC   the activity of PKH by expressing inactive proteins or to supplement the
CC   patients own production of PKH polypeptides. Additionally, the DNA may be
CC   used to produce PKH, according to standard recombinant DNA methodology,

```

CC by inserting the nucleic acids into a host cell and culturing the cell to
 CC express the protein. Conversely, antisense nucleic acid molecules may be
 CC administered to down regulate PKH expression by binding with the cells
 CC own PKH genes and preventing their expression. The DNA, and antisense
 CC sequences may also be used as DNA probes in diagnostic assays to detect
 CC and quantitate the presence of similar nucleic acid sequences in samples,
 CC and hence which patients may be in need of restorative therapy. They may
 CC also be used to study the expression and function of PKH polypeptides and
 CC their role in metabolism. The PKH polypeptides may be used as antigens in
 CC the production of antibodies against PKH and in assays to identify
 CC modulators (agonists and antagonists) of PKH expression and activity. The
 CC anti-PKH antibodies and PKH antagonists may also be used to down regulate
 CC PKH expression and activity. The anti-PKH antibodies may also be used as
 CC diagnostic agents for detecting the presence of PKH polypeptides in
 CC samples.

XX Sequence 1846 BP; 611 A; 321 C; 369 G; 545 T; 0 other;

Alignment scores:

Quality: 1959.00 Length: 438
 Ratio: 4.922 Gaps: 2
 Percent Similarity: 90.868 Percent Identity: 89.269

Alignment block:

US-09-783-320-4 x AA286797 ..

Align seg 1/1 to: AA286797 from: 1 to: 1846

```

1 MetGluTyrValArgLeuGlnLysIleGlyGluGlySerPheGly 17
154 ATGGAGAAAGTATGTTAGACTACAGAAAGATTGGAGAGGTTCAATTTGGAAA 203
17 sAlaIleuValLysSerThrGluAspGlyArgGlnTyrValIleLysG 34
204 AGCCATCTCTGTTAATCTACAGAAAGATGGCAGACGATGTTATCAAG 253
34 IuileSnIleSerArgMetSerSerLysGluArgGluGluSerArgArg 50
254 AATTTAAGCTCTCAAGATGTCAGTAAAGAAAGAAAGAAATCAAGAGA 303
51 GluValAlaValIleuValAsnMetLysHisProsnIleValGlnTyr 67
304 GAAATGCGAGTATTTGGCAACATGAGCATCCAAATATGTCAGATAG 353
67 gGluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
354 AGAATCATTTTGA..... 366
84 IuileGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
366 ..... 366
101 GlnGluAspGlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeu 117
367 .....GGAATTTTGGACTGCTTTGTACAGATATGTTGGCCCTGAA 407
117 sHisValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsn 134
408 ACAATGTCATGATAGAAAATTCCTCATCGAGACATTTAAATCTCAGACA 457
134 IePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
458 TATTTTAACTAAAGATGAGACAGTACACCTTGAGATTTTGGAAATGCT 507
151 ArgValLeuAsnSerThrValGluLeuAlaArgThrCysIleGlyThrPr 167
508 AGACTTCTTAATAGTACTAGAGCGGCTCGAACTTCATAGGGAGCC 557
167 cTyrTyrLeuSerProGluIleCysGluAsnLysProTyrAsnAsnLys 184
558 ATACTACTGTCACTGAATCTGTGAACAAACAACTTACATATATAAA 607

```

```

184 eAspIleTrpAlaLeuGlyCysValLeuTyrGluLeuCysThrLeuLys 200
608 GTGACATTTGGCGCTGGGGGTGTCTCTTATGACCTGTACACTTAA 657
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValIleLysIleLeu 217
658 CATGCTTTGAGCTGACATATGAAAACCGTACAGAAAGATTAATATC 707
217 rGlySerPheProValSerLeuHisTyrSerTyrAsnLeuArgSerL 234
708 TGGATCTTTTCACCTGTGTCTTGTGCAATTTCTCATGATCTCCCACT 757
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
758 TGGTGTCTCACTGTATTTTAAAGAAATCTAGGATAGACATCACTGAC 807
251 SerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPheLeu 267
808 TCCATATTTGAGAAAGGTTTATATAGCCAAACGCAATTTTCTCTC 857
267 rProGlnLeuIleAlaGluGluPheCysLeuLysThrPheSerLysPheG 284
858 TCCTCAGCTATTTGCAGAGCAATTTGTCTTAAACATTTTGCAGATTG 907
284 lYserGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
908 GATCACACCTATATCCACCTAAAGACACCTTCAGAGCAAAACCTGANT 957
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGly 317
958 TCTGTTATGCTGCTCAAAATTTCAAAAGCTGCGCTAAATATGANT 1007
317 eProLeuAlaTyrLysLysTyrGlyAspLysValLeuHisGlyLysP 334
1008 ACCTTTAGCATATTAAGAAATTTGAGATTAATAATTTACAGAAAGAAC 1057
334 rLeuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
1058 CACTGCAAAACATTAACAGGCCCATCAACTCCAGAGAAAGAGATGANT 1107
351 ThrGlyGluGluArgArgLysIleSerGluGlnAlaAlaArgLysArg 367
1108 ACTGGAAGAAAGAGAGAAATATCTAGAGACAGCAAGAAAGAGAG 1157
367 gLeuGluPheIleGluLysGlyLysLysGlnLysAspGlnIleLeuSer 384
1158 GCTGGAATTTATTTGAAAAGATTAAGAACGGTAGGATCAGATTATTA 1207
384 euMetLysAlaGluGlnMetLysArgGlnGlnLysGluArgLeuGluArg 400
1208 TAATGAAGGCTGAACAAATGAAGCAAGCAAGCAAGCAAGCTTGCAAGA 1257
401 IleAsnArgAlaArgGluGlnGlyTrpArgAsnValIleuSerAlaGly 417
1258 ATTAATATGCGCCAGCGAACAAGATGAGAAATATGCTAAAGTGTG 1307
417 ySerGlyGluValLysAlaProPheLeuGlySerCysLys.....Thr 432
1308 AAGTGGTGAAGTAAAGTATTTATACCAATATATGTTATATCTACCA 1357
432 IeAlaProSerSer 436
1358 TTTTCCCTCCAGT 1371

```

seq_name: /nef/abs06/SIDSL/9cgcdata/hold_geneseg/genesegn-emb1/NA2001A.DAT:AA011848

seq_documentation_block:

ID AAD11848 standard; CDNA: 1846 BP.

XX AAD11848;

AC

XX

DT 25-SEP-2001 (first entry)

XX

DE Human protein kinase homolog-6 (PKH-6) cDNA.

XX Human; protein kinase homolog-6; PKH-6; cytostatic; protein therapy;
 KW vaccine; immunosuppressive; antisclerotic; antiabortive; adenocarcinoma;
 KW Acquired immune deficiency syndrome; AIDS; melanoma; cancer; bone; liver;
 KW breast; autoimmune disorder; multiple sclerosis; drug screening; anaemia;
 KW Crohn's disease; ectopic pregnancy; tubal disease; inflammatory disorder;
 KW reproductive disorder; polycystic ovary syndrome; asthma; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 154..1191
 FT /tag-#
 FT /product="Human protein kinase homolog-6 (PKH-6)"

XX US6264947-B1.

XX 24-JUL-2001.

XX 20-OCT-1999; 9905-0420915.

XX 15-OCT-1998; 9805-0173581.

XX (INCYTE GENOMICS INC.

XX Bandman O, Tang YF, Hillman JL, Yue H, Guegler KJ, Corley NC;
 PI Gorgone GA, Azimzai Y, Lu DAM;
 DR WPI: 2001-450728/48.
 DR P-PSDB: AAE06211.

PT Human protein kinase proteins and homologs, useful for preventing,
 PT diagnosing and treating cancers, autoimmune/inflammatory disorders and
 PT reproductive disorders.

XX Example; Column 67-68; 38pp; English.

XX The present cDNA sequence encodes human protein kinase homolog-6 (PKH-6).
 CC Human protein kinase homologs (PKH) and their cDNA molecules are used in
 CC the prevention, diagnosis and treatment of diseases associated with
 CC increased or decreased expression of PKH. Examples of such disorders
 CC include, cancer (e.g. adenocarcinoma, melanoma and bone, breast and
 CC liver cancer), autoimmune/inflammatory disorders (e.g. Acquired Immune
 CC deficiency Syndrome (AIDS), anaemia, asthma, Crohn's disease and
 CC multiple sclerosis) and reproductive disorders (e.g. tubal disease,
 CC ectopic pregnancy and polycystic ovary syndrome). PKH, its catalytic or
 CC immunogenic fragment are used for screening libraries of compounds in any
 CC of the drug screening techniques. PKH nucleic acids are used to generate
 CC hybridisation probes useful in mapping the naturally occurring genomic
 CC sequences. PKH are also used as antigens in the production of antibodies
 CC against protein kinases (PK) and in assays to identify modulators of PK
 CC expression and activity. PKH is also used in protein therapy.

XX Sequence 1846 BP; 611 A; 321 C; 369 G; 545 T; 0 other:

XX alignment_scores:

XX Quality: 1959.00 Length: 438
 XX Ratio: 4.922 Gaps: 2
 XX Percent Similarity: 90.868 Percent Identity: 89.269

XX alignment_block:

XX US-09-783-320-4 x AAD11848

XX Align seg 1/1 to: AAD11848 from: 1 to: 1846

1 MetGluTyrValArgLeuGlnLysIleGlyLysPheGlyIle 17
 154 ATGGAGAGTAGTGTAGACTACAGAGATTGACAGACGTTCAATTGGAAA 203
 17 sallaileuvallyserthrighuaspglyargintyrvalilleysg 34
 ||||||||||||||||||||||||||||||||||||||||||||

204 AGCCATTCCTGTTAAATCTACAGAGATGCGACAGATGTTATTCACAG 253
 34 IuileasnlIeserArgMetSerSerLysGluArgGluLysSerArg 50
 254 AAATTAACATCTCAAGAAATGTCCAGTAAAGAAAGACAAATCAAGAGGA 303
 51 GluValAlaValLeuAlaAsnMetLysHisProAsnIleValGlnTyrAr 67
 304 GAAGTTCACGATTTGGCAACATGAAGCATCCCAATATTTGTCACGATATG 353
 67 gGluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
 354 AGAATCATTTGAA..... 366
 84 LuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
 366 366
 101 GlnGluAspGlnIleLeuAspTrpPheValGlnIleCysLeuAlaLeuL 117
 367GCAATTTTGACTGTTTGTACAGATATGTTTGCCCTGAA 407
 117 SHsValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsnI 134
 408 ACATGTACATGATAGAAAATTTCTCATCGACATTAATCATCAGAACAA 457
 134 IephLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
 458 TATTTTAACATAAAGATGGAACAGTACAACCTTGAGATTTTGGAAATGCT 507
 151 ArgValLeuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr 167
 508 AGAGTTCTTAATAGTACTAGTAGAGCTGCTCGAAGCTTGATAGGACACC 557
 167 cTyrTyrLeuSerProGluIleCysGluAsnLysProCysAsnAsnLys 184
 558 ATACTACTGTGACCTGAAATCTGTGAACAAACCTTACAAATATATAA 607
 184 eRaSpIleTrpAlaLeuGlyCysValLeuTyrGlnLeuCysThrLeuLys 200
 608 GTGACATTTGGCGCTGTGGGTGTGCTTTATCAGCTGTGTAACCTTAA 657
 201 HisAlaPheGlnAlaGlySerMetLysAsnLeuValLeuLysIleLeSe 217
 658 CATGCTTTTGAAGCTGCGACATATGAAAACCTGGTACTGAAGATATATTC 707
 217 rGlySerPheProProValSerLeuHisTyrSerTyrAspLeuArgSerL 234
 708 TGGATCTTTTCCACCTGTGCTTGCATATATTCCTATGATTCCTCCAGTT 757
 234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
 758 TGGTGCTCAGATTATTTAAAGAAATCTTAGGGATAGACCATCAGTCAAC 807
 251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGlnLysPheLeuSe 267
 808 TCCATATTGGAGAAAGGTTTATATAGCCAAACGATTTGAAAATTTCTCTC 857
 267 rProGlnLeuIleAlaGlnLysPheCysLeuLysThrPheSerLysPheG 284
 858 TCCTCACCTTATTCAGAAAGAAATTTTGTCTTAAAAACATTTTGAAGTTG 907
 284 LysSerGlnProIleProAlaLysArgProAlaSerGlyLysAsnSerIle 300
 908 GATCAGACCTATACAGCTTAAAGAACAGCTTCAGAGCAAAACATCGATT 957
 301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGlyI 317
 958 TCTGTATGCTGCTGCGAAGAAATTTACAAAGCCGCGCGCTAATATGGAAT 1007
 317 eProLeuAlaTyrLysLysTyrGlnLysAspLysLeuHisGlnLysLysP 334
 1008 ACCTTTAGCATATTAAGAAATATGAGATATAAAATTTACACGAAAGAAAC 1057

```

334  ToleuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
1058  CACGCGCAAAAAACATTAACAGCCCAACATCAACAGAGAGAGAGAGAT 1107
351  ThrGlyLysGlnLysArgLysLysSerGluGlnAlaAlaLysArgVal 367
1108  ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1157
367  GLeuGlnPheLysGlnLysGlnLysLysSerGlnLysSerGlnLys 384
1158  GCTGGAATTTATTTGAAAAGATTAAGAGAGAGAGAGAGAGAGAGAG 1207
384  eumetLysAlaGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 400
1208  TATGTAAGCTGACAAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
401  IleAsnArgAlaArgGlnGlnLysGlnLysGlnLysGlnLysGlnLys 417
1258  ATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1307
417  ySerGlyLysValLysAlaPheLysGlnLysGlnLysGlnLysGlnLys 432
1308  AAGTGTGTAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
432  LeAlaProSerSer 436
1358  TTTTCCCTCCAGT 1371

seq_name: /net/abs06/sids1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAH06097

seq_documentation_block:
ID  AAH06097 standard; cDNA; 876 BP.
AC  AAH06097;
DT  26-JUN-2001 (first entry)
DE  Human cDNA clone (5'-primer) SEQ ID NO:2932.
KW  Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
OS  Homo sapiens.
PN  EPI074617-A2.
PD  07-FEB-2001.
PF  28-JUL-2000; 2000EP-0116126.
PR  29-JUL-1999; 99JP-0248036.
PR  27-AUG-1999; 99JP-0300253.
PR  11-JAN-2000; 2000JP-0118776.
PR  02-MAY-2000; 2000JP-0183767.
PR  09-JUN-2000; 2000JP-0241899.
XX  (HELI-) HELIX RES INST.
XX  Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX  Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX  WPI: 2001-318749/34.
XX  Primer sets for synthesizing polynucleotides, particularly the 5602
XX  full-length cDNAs defined in the specification, and for the detection
XX  and/or diagnosis of the abnormality of the proteins encoded by the
XX  full-length cDNAs -
XX  Claim 1: SEQ ID 2932: 2537bp + CD ROM; English.
XX  The present invention describes primer sets for synthesizing 5602
XX  full-length cDNAs defined in the specification. Where a primer set
XX  comprises: (a) an oligo-dT primer and an oligonucleotide complementary

```

```

CC  to the complementary strand of a polynucleotide which comprises one of
CC  the 5602 nucleotide sequences defined in the specification, where the
CC  oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC  of an oligonucleotide comprising a sequence complementary to the
CC  complementary strand of a polynucleotide which comprises a 5'-end
CC  sequence and an oligonucleotide comprising a sequence complementary to a
CC  polynucleotide which comprises a 3'-end sequence, where the
CC  oligonucleotide comprises at least 15 nucleotides and the combination of
CC  the 5'-end sequence/3'-end sequence is selected from those defined in
CC  the specification. The primer sets can be used in antisense therapy and
CC  in gene therapy. The primers are useful for synthesizing polynucleotides,
CC  particularly full-length cDNAs. The primers are also useful for the
CC  detection and/or diagnosis of the abnormality of the proteins encoded by
CC  the full-length cDNAs. The primers allow obtaining of the full-length
CC  cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC  AAH13633 to AAH18742 represent human cDNA sequences. AAH92446 to
CC  AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC  represent oligonucleotides, all of which are used in the exemplification
CC  of the present invention.
XX  SQ
Sequence 876 BP: 307 A; 160 C; 159 G; 239 T; 11 other:

alignment_scores:
Quality: 1218.50      Length: 282
Ratio: 4.598          Gaps: 5
Percent Similarity: 93.972 Percent Identity: 89.716

alignment_block:
US-09-783-320-4 x AAH06097
Align seg 1/1 to: AAH06097 from: 1 to: 876

91  ArgIleAsnAlaGlnLysGlyValLeuPheGlnLysSerGlnLysLeuAsn 107
|||||
3  CGAATAAATGCTCAGAAAGCGCTTTGTTTCAGAGATCAGATTGTGGA 52
107  pTrpPheValGlnLysCysLeuAlaLeuLysHisValHisAspArgLys 124
|||||
53  CTGCTTTGACAGATATGTTTGCCCTGAAACATGATACATGAGAAAAA 102
124  LeuLysHisArgAspLysSerGlnAsnLysPheLeuThrLysAspGly 140
|||||
103  TTCTTCATCGACAGCATTAATCTCAGAAACATATTTTAACTTAAGATGGA 152
141  ThrValGlnLeuGlyAspPheGlyIleAlaArgValLeuAsnSerThrVal 157
|||||
153  ACAATCAACACTTGAGATTGGAATGCTAGAGTCTTAATTAAGTACTGT 202
157  IGluLeuAlaArgThrCysIleGlyThrProTyrTyrLeuSerProGlu 174
|||||
203  AGAGCTGGCTCGAACTTGATAGGAGCCCACTACTGTCACCTGAAA 252
174  LeuGlyAsnLysProTyrAsnAsnLysSerAspIleTyrAlaLeuGly 190
|||||
253  TCTGTGAAAAAACAACCTTAATTAATTAATTAATTAATTAATTAATTA 302
191  CysValLeuTyrGlnLeuCysThrLeuLysHisAlaPheGlnAlaGly 207
|||||
303  TGATGCTTTATGAGCTGTATACACTTAACATGCTTTTGAAGCTGGCAG 352
207  TmetLysAsnLeuValLeuLysIleLeuSerGlySerPheProProVal 224
|||||
353  TATATAAAACCTGCTACTGAAATATATATCTGATCTTTTCCACCTGCT 402
224  erLeuHisLysSerTyrAspLeuArgSerLeuValSerGlnLeuPheLys 240
|||||
403  CTTTGCAATTAATCTATGATCTCCGAGTTGGTGGTACGATTAATTTAA 452
241  ArgAsnProArgAspArgProSerValAsnSerIleLeuGluLysGlyPhe 257
|||||
453  AGAATATCTAGGAGATGAGCATTGATCACTCAATTAATTTGAGAAAGTTT 502

```

```

seq_documentation_block:
ID_AAS06745 standard; cDNA; 1947 BP.
XX
AC_AAS06745;
XX
DT_12-SEP-2001 (first entry)
XX
DE_Polynucleotide sequence encoding human protein kinase #45.
XX
XX_Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
XX_KM metabolic disorder; immune related disease; neurological disorder;
XX_KM neurodegenerative disorder; inflammatory disorder; infectious disease;
XX_KM reproductive disorder; gene therapy; ss.
XX
OS_Homo sapiens.
XX
XX_WO200138503-A2.
XX
PD_31-MAY-2001.
XX
PF_22-NOV-2000; 2000MO-US32085.
XX
XX_24-NOV-1999; 99US-0167482.
XX
XX_(SUGEN-) SUGEN INC.
XX
XX_Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
XX_PI Flanagan P, Clary D;
XX_PI
XX_WPI: 2001-343950/36.
XX_DR P-PSDB; AAU03545.
XX
PT_Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX
XX
PS_Example 1; Figure 1; 433pp; English.
XX
XX_AAS06701:AAS06757 encode for novel human protein kinases #1-57. The
XX novel protein kinases have been identified as members of the tyrosine
XX or serine/threonine kinase (PTK and STK) families. The polynucleotides
XX encoding protein kinases and the polypeptides may be used in the
XX prevention, diagnosis and treatment of diseases associated with
XX

```

```

alignment_scores:
  Quality: 1025.00      Length: 818
  Ratio: 2.293          Gaps: 25
  Percent Similarity: 54.645      Percent Identity: 32.763

alignment block:
  US-09-783-320-4 x AAS06745      ..

Align seg 1/1      to: AAS06745      from: 1      to: 1947

```

[illegible]

```
200 LysHisIalaphenIuaIagIyserMetLysAsnLeuValleuIleI 216
|||||
601 AAMCARTCTTTTGAGGGTAACTTACACACCTGTTCTGAAGATTG 650
|||||
216 eSerGlySerPheProValSerLeuHisTyrSerTyrAspLeuArgS 233
|||||
651 TCACACACATTTTGGCCCCAATATCTCCGGGGTTTTCTCGAGACTTCAT 700
|||||
233 eLeuValSerGlnLeuPheLysArgAsnProAlaArgAsnProSerVal 249
|||||
701 CCTGATATCTCAGCTCTTCAAGTATCTCTGAGACCGACATCATATA 750
|||||
250 AsnSerIleLeuGlnLysGlyPheIleAlaLysArgIleGlnLysPhe 266
|||||
751 AATTCATTTTCAAAAGGCCCTTTTAGAGATCTATTCCAAAATATT 800
|||||
266 uSerProGln.....LeuIleAlaGlnLup 275
|||||
801 GACTCTGAGGTAACTTTGAGGTGACTGTTGATTTTGCACAGATTT 850
|||||
275 heCysLeuLysThrPheSerLys.....PheGlySerGlnProIle 288
|||||
851 TGGGTTCAGGTCCTTGACAGCTGTGTCGGTTTACGTGTGAGCCACTG 900
|||||
289 ...ProAla.....LysArgProAlaSerGlnAsnSe 299
|||||
901 CACCCAGCCTGTATTATGTTTTTAAACATCCCTCTGTTCTTCTTCA 950
|||||
299 rIleSerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrG 316
|||||
951 GATATAAATGATAGAAAGACCACCAAAATGCTGCTGTGCGACATATG 1000
|||||
316 LylLeuProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGlnLys 332
|||||
1001 ATTATTATTATGCTCAACTTGATATGCTGAGAGAGAGGCCAC..... 1044
|||||
333 LysProLeuGlnLysHisLysGlnIleHisGlnThrProGlnLysArgVa 349
|||||
1045 AAMCCA.....AGTTATCACCCCTATTCCTCAAGAA..... 1074
|||||
349 LAsnThrGlyGlnLysLysArgLysIleSerGlnLysAlaIleArgLysA 366
|||||
1075 .AATCTGAGAGTTGAG..... 1089
|||||
366 rGArgLeuGlnPheIleGlnLysGlnLysLysGlnLysAspGlnIleIle 382
|||||
1090 .....GATTACGGCTCAGAAACGAGCATGGTCCATCCCA..... 1125
|||||
383 SerLeuMetLysAlaGlnLysMetLysArgGlnGlnLysGlnArgLeuG 399
|||||
1126 AGTCATATGGCTGCTGAGTACCTTCAAGAGAAATTGGAAGCTCAACAATA 1175
|||||
399 uArgIleAsnArgAlaArgLysGlnGlnLysTyrPArgAsnValLeuSerAlaG 416
|||||
1176 TAACTTGAAAGTGAGAAACAAATTTGGTCTGCT..... 1209
|||||
416 LysGlySerGlyLysValLysAlaPheLeuGlnLysSerGlyLysThrIle 432
|||||
1209 ..... 1209
|||||
433 AlaProSerSerPheSerSerArgGlyGlnTyrGlnHisTyrHisAlaI 449
|||||
1210 ...CCATCTTCTGCCGAGCCAAT..... 1230
|||||
449 ePheAspGlnMetGlnGlnGlnIleArgAlaIuaAspAsnGlnIuaIaLysTrpL 466
|||||
1231 .TACACACAGACACAGAGCTAAGAGTATGAGAGAAAGCCTAGATTCC 1279
|||||
466 LysArgGlnIleTyrGlyLysArgLysLeuProGlnLysGlnLysGlnLeu 482
|||||
1280 AGGAG.....CTGCCATTTTGGAAAAAACCAATGAAAG 1311
|||||
483 AlaValGluArgAlaLysGlnValGlnLysPheLeuGlnArgLysArgL 499
|||||
1312 GAACAGCAAAATATGCAACAGCACTTAGAGAA..... 1341
|||||
499 uAlaMetGlnAsnLysAlaArgAlaGlnGlnHisMetValTyrLeuAlaA 516
|||||
1341 ..... 1341
|||||
516 rGLeuArgGlnIleArgLeuGlnAsnPheAsnGlnArgGlnGlnIleLys 532
|||||
1342 .....ATAGCCCAACAGTACCAATGACATGACATGAAAGAAATTAAG 1380
|||||
533 AlaLysLeuArgGlyGlnLysLysGlnAlaAsnHisSerGlnLysGlnG 549
|||||
1381 AACAGATGGGAGAGAAACAGAGGACTT..... 1410
|||||
549 uGlySerGlnGlnAlaAspMetArgArgLysLysIleGlnSerLeuLysA 566
|||||
1411 .....GAAAAAGACTTGAAACAAATGAGGCTTCAGAACACAAAG. 1449
|||||
566 IahIslaAsnAlaArgAlaAlaValleuLysGlnGlnLeuGlnLys 582
|||||
1450 .....GAAAGTAA 1458
|||||
583 ArgLysGlnAlaTyrGlnArgGlnLysLysValTrpGlnGlnHisLeuVa 599
|||||
1459 AATCCAGAA.....CAGAAATTTAAAGC 1481
|||||
599 LAlaLysGlyValLysSerSerAspValSerProLeuGlnHisG 616
|||||
1482 TAAAGAGGGGTAAATT..... 1500
|||||
616 LuthrGlyGlySerProSerLysGlnGlnMetArgSerValIleSerVal 632
|||||
1500 ..... 1500
|||||
633 ThrSerAlaLeuLysGlnValGlyValAspSerSerLeuThrPheThrAr 649
|||||
1501 .....GAAATATTTAGCAAAATGATTTCTGAT..... 1530
|||||
649 gGlnThrSerGlnLysMetGlnLysThrAsnAlaIleSerSerLysA 666
|||||
1530 ..... 1530
|||||
666 rGlnIleLeuArgArgLeuAsnGlnAsnLeuLysAlaGlnGlnAspGln 682
|||||
1531 .....GAAACATCTCTCCAAAGAGAGAGCA 1557
|||||
683 LysGlyMetGlnAsnLeuSerAspThrPheGln.....IleAs 695
|||||
1558 ATGGATATACCAAAATGAACCTTGACCTTTCAGAGATGCGATGAATTTAA 1607
|||||
695 nValHisGlnAspAlaLysGlnHis.....G 704
|||||
1608 GGAATATGAATGCTAAGAGCATGAGATTTACAGACAAAGCATTTTG 1657
|||||
704 Lulys.....GlnLysSerValSerSer..... 711
|||||
1658 AAAAATCTCACTGCCCAAGAGAGGGTTTTTCCAGCAGACTGTAGCTGCT 1707
|||||
712 .....AspArgLysLysTrpGlnAlaGlyLys...GlnLeuValIlePr 725
|||||
1708 GTGGCAACAGAGAGAGCACTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 1757
|||||
725 oLeuAspGlnLeuThrLeuAspThrSerPheSerThrThrGlnAlaGlnHisT 742
|||||
1758 GATGATGGCAGCTGGCGACATCACTTCACCTGACCTCCCAAGGGGCGTGACA 1807
|||||
742 hrValGlyGlnValIleLysLeuGlnLysProAsnGlySerProArgAla 758
|||||
1808 AT...GGCCAAAGTTATTTGATTAAGGATTCAGAGAAACAGAGAAACAG 1854
|||||
759 TrpGlyLysSerProThrAspSerValLeuLysIleLeuGlyGlnAlaG 775
```


• Wed May 15 14:14:00 2002

649 rggju 650
|||||
848 GCGMA 852

us-09-783-320-4.rng

THIS PAGE BLANK (USPTO)

Copyright (c) 1993-2000 CompuGen Ltd

-NO_XLPRXY -WAIT -THREADS=1

Escrow Fee

gb_ro:BC010302 + 556.00 519.27 1.5e-20 3100

2025-01-01 10:00:00

3645 bp DNA linear PAT 10-SEP-2001

First of all, we have to find the value of α and β such that

FIELD; PLINIAES, CACALINUT, MONITORING, 3645)

and polynucleotides encoding the same

tics Incorporated (US)

.3645

```

_xref1="taxon:9606"
003 +
004 +

```

Length: 1214

Percent Identity: 100.000

14-00000 1 401 364E

17

AGACTACAGAAGATTGGAGAAGGTTTCATTGGAAA 50

```
ssErthrgluaspGlyArgGlnTyvalIleLysG 34
```

ATCTACAGAGA TGGCAGACAGTATGTATCAAGG TTT

argmetsetserlysluajguuuberaiyuaig

THE UNIVERSITY OF CHICAGO

[illegible]

best was 100% at 100

TTTAACTCAATAAATCTTCAGAAAGCGTTTGT 300

eLeuAspTrpPheValGlnIleCysLeuAlaLeuLy 117

TTTGGACTGGTTTGTACAGATATGTTTGGCCCTGAA 350

rglysiLeuNHsArprilelysselglnsh 134

GAAAAATCTCAGCCGACCAATTTCCTGGTAACTTT

```
134 IepheLeuThrLysasp61yThrValGlnLeuGlyaspPheGlyIleAla 150
|||||
401 TATTTTAACTAAAGATGCAACACTGACACTTGGAGATTTTGGAAATGGCT 450
|||||
151 ArgValLeuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr 167
|||||
451 AGAGTTCTTAATAGTACTGTAGACTGCTGCAGAACTGCTAGAGGAGCC 500
|||||
167 CTYTYrLeuSerProGluIleCysGluAsnLysProTYrAsnAsnLys 184
|||||
501 ATACTACTGTCCACTGAAATCTGTGAAACCAACCTTACATAATAAATA 550
|||||
184 eAspIleThrAlaLeuGlyCysValLeuTYrGlnLeuGlySThrLeuLys 200
|||||
551 GTGACATTTGGGCTCTGGGGTGTCTCTTATGAGCTGTGTACACTTAA 600
|||||
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleIleSe 217
|||||
601 CATGCTTTTGAAGCTGGCAGATGAAACCTGTACTGTAAGATTAATATC 650
|||||
217 rGlySerPheProValSerLeuHisTYrSerTYrAspLeuArgSerL 234
|||||
651 TGGATCTTTTCCACCTGTCTCTTGCATTTATCTCATGATCTCCGAGTT 700
|||||
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
|||||
701 TGGTCTCTCACTTATTTAAAGAAATCTAGGAGTAGACCATCAGTCAAC 750
|||||
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGlnLysPheLeuSe 267
|||||
751 TCCATATTTGGAGAAAGTTTATAGCCAAACGATTGCAAAAGTTTCTCC 800
|||||
267 rProGlnLeuIleAlaGlnLysPheCysLeuLysThrPheSerLysPheG 284
|||||
801 TCTCTAGCTTATTCAGAGAAATTTTCTTAAACATTTTCCAAAGTTTG 850
|||||
284 LysSerGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
|||||
851 GATTCACGACCTATACACAGCTTAAAGACCCCTTACAGCAAAACCTGAT 900
|||||
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTYrGlyI 317
|||||
901 TCTGTATAGCCCTGCTCAGAAATTAACAACCTGCCCTTAATATGGAAT 950
|||||
317 eProLeuAlaTYrLysLysTYrGlyAspLysLysLeuHisGlnLysLysP 334
|||||
951 ACCTTTACCATATTAAGAAATATGAGAGATTAATAAATTTACACGAAAGAAC 1000
|||||
334 rOLEuGlnLysHisLysGlnAlaHisGlnThrProGlnLysArgValAsn 350
|||||
1001 CACTCCAAAAACATTAACAGGCCCATCAACCTCCAGAGAAAGAGATGAAT 1050
|||||
351 ThrGlyGlnArgArgLysIleSerGlnGlnAlaAlaArgLysArgArg 367
|||||
1051 ACTGAGAAAGAAAGAGAGAAATATCTGAGSAAAGCAAGCAAGAAAGAGAG 1100
|||||
367 GlnLeuLysPheIleGlnLysGlnLysLysGlnLysAspGlnIleIleSerL 384
|||||
1101 GCTGGAATTTATGAAAGAAAGAAAGAAAGATCAGATTAATTAATGAT 1150
|||||
384 euMetLysAlaGlnLysMetLysArgGlnLysGlnLysGlnLeuGlnArg 400
|||||
1151 TAAATGAAGCTGAACAAATGAAGCAAGCAAGCAAGCAAGCAAGCTTGGAAG 1200
|||||
401 IleAsnArgAlaArgGlnGlnIYrPArgAsnValLeuSerAlaGlyI 417
|||||
1201 ATAAATAGCGCCAGGAAAGCAAGATGAGAAATGTCTAGTCTGCTGCTG 1250
|||||
417 ySerGlyGlnValLysAlaProPheLeuGlnLysGlyIleTYrHisIleAlaP 434
|||||
1251 AAGTGTAAGATTAAGGCTCTTTCTTGGGAGAGTGGAGGACATATAGCTC 1300
|||||
434 rSerSerPheSerSerArgGlyGlnIYrGlnHisTYrHisAlaIlePhe 450
|||||
1301 CATCATCTTTTCTCTCCGAGCAGACTATGACAACTTACCATGCCATTTT 1350
|||||
451 AspIleMetGlnGlnIleArgAlaGluAsnAsnGlnAlaLysTYrPLeuArg 467
|||||
1351 GACCAATATGACAGCAACAAAGAGAGAAATTAATGAAGCTTAATGGAAAG 1400
|||||
467 ggLysIleTYrGlyArgGlyLeuProGlnArgGlnLysGlnLysAlaVal 484
|||||
1401 AGAAATATATGCTCGAGGCTTCTCCAAAGCAAGCAAGCAAGCAAGCTAG 1450
|||||
484 alGlnArgAlaLysGlnValGlnLysPheLeuGlnArgGlyArgGlnAla 500
|||||
1451 TAGAAAGAGCTTAAACAGTAGAAGAGCTTCTCAGAGCAAGAAAGGAGAGCT 1500
|||||
501 MetGlnAsnLysAlaArgAlaGlnLysMetValTYrLeuAlaArgLe 517
|||||
1501 ATGCAAGATTAAGCTCGAGCCGAGAGACATATGATTTATCTGGCAAGACT 1550
|||||
517 uArgGlnIleArgLeuGlnAsnPheAsnGlnArgGlnGlnIleLysAlaL 534
|||||
1551 GAGCAAAATTAAGACTACAGAAATTTCAATGAGCGCCACAGATTAAGCCA 1600
|||||
534 ySLeuArgGlyGlnLysLysGlnAlaAsnHisSerGlnGlyGlnLysL 550
|||||
1601 AACTTGTGTGAAAGAAAGAAAGCTAATCATTTCTGAAGAGCAAGAAAGCA 1650
|||||
551 SerGlnGlnAlaAspMetArgArgLysIleGlnSerLeuLysAlaHis 567
|||||
1651 AGTGAAGAGCTGACATGAGGGCGCAAAAATCGAATCATCTGAAGCCCA 1700
|||||
567 sAlaAsnAlaArgAlaAlaValLeuLysGlnGlnLeuGlnArgLysArgL 584
|||||
1701 TGCAAAAGCAGCTGCTGCTGTACTAAAGAAACAATAGACCAAGCAAGAA 1750
|||||
584 ySGLuAlaTYrGlnArgGlnLysValTYrGlnGlnHisLeuValAla 600
|||||
1751 AGGAGGCTTATGAGAGAGAAAGAAAGTGTGGAAAGCATTTGGTGCT 1800
|||||
601 LysGlnValLysSerSerAspValSerProProLeuGlnHisGluTh 617
|||||
1801 AAAGGAGTTAAGAGTTCTGATGTTTCTCCACCTTTGGGACACCAAGAAC 1850
|||||
617 rGlyLysSerProSerLysGlnLysMetArgSerValIleSerValIThrS 634
|||||
1851 AGTGCGCTCTCCATCAACCAACAGATGAGATCTGTATTCTGTAACTT 1900
|||||
634 eAlaLeuLysGlnValGlyValAspSerSerLeuThrAspThrArgGlu 650
|||||
1901 CAGCTTGAAGAAAGTTGGGGTGAGCAGTAGTTTAACTGATACCGGGGAA 1950
|||||
651 ThrSerGlnLysMetGlnLysThrAsnAsnAlaIleSerSerLysArgG 667
|||||
1951 ACTTCAGAAAGAGATGCAAAAAGACCAACATGCTATTCTCAAGTAAGCGGA 2000
|||||
667 uIleLeuArgArgLeuAsnGlnLysAsnLeuLysAlaGlnLysAspGlnLysG 684
|||||
2001 AATATCTCGCAGATTAATGAATAATCTTAAGCTCAAGAGAGTGAAGAAAG 2050
|||||
684 LysMetGlnAsnLeuSerAspThrPheGlnIleAsnValHisGlnAspAla 700
|||||
2051 GAATGCGCAATTCCTGTACTTTGAGATTAAGATTAAGTTCAATCAATCAAT 2100
|||||
701 LysGlnHisGlnLysGlnLysSerValIleSerSerAspArgLysTYrP 717
|||||
2101 AAAGAGCATGAAGAAAGAAAGAAATCAGTTTCATCTGATCCGAGAAATGGGA 2150
|||||
717 uAlaGlyGlnLeuValIleProLeuAspGlnLeuThrLeuAspThrS 734
|||||
2151 GCGAGAGAGTCAACTTGTGATCTCTCTGATGAGTTAAGTAAACATGAATAT 2200
|||||
734 eRPheSerThrThrGlnArgHisThrValGlyGlnValIleLysLeuLys 750
|||||
```

2201 CCTCTCTACACTGAAAGACATACAGTGGAGAAATTATTAAATTAGT 2250
 751 ProasnglyserProargargalatrpglylyserProthrAspSerVa 767
 2251 CCTAATGATCTCCAAAGAGAGCTGGGGAAAGTCCGACAGATTCTGT 2300
 767 lLeuLyllaleuGlulaleuGlulaleuGlulaleuGlulaleuGlul 784
 2301 TCTAAATATCTGGAGAAAGCTGACACTGACAGACTGACAGACTATTA 2350
 784 lAspThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 800
 2351 AAATACACTATTAGAGAGATTTTCCCGAAGGGGAAAGATGACAA 2400
 801 ProLeuThrThrThrThrThrThrThrThrThrThrThrThrThrThr 817
 2401 CCTTAATCTCTGGAGAAAGAGTCAATGTTTTCATGCAATGAAATTA 2450
 817 nProSerAlaIleValAspSerProValGluThrThrThrThrThrThr 834
 2451 CCCATCAGCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2500
 834 eArgLualSerProGlnMetSerLeuLysLeuGlnLysLeuGlnLys 850
 2501 GTAGAGCATCTCCACAGATGTCATGAACTGGAGAGAAATTTAGAGA 2550
 851 ProAspAspLeuGluThrGluThrGluThrGluThrGluThrGluThr 867
 2551 CCTGATGATTGGAAACGAAATCTCAAGAGCCCAAGTGGAAACAA 2600
 867 sAspGlySerLeuProCysThrIleThrAspValTyrPileSerGluL 884
 2601 AATGAGAGCTTGGCCATGCTGATGATGATGATGATGATGATGATG 2650
 884 ySgLTThrLysGluThrGlnSerAlaAspArgIleThrIleGlnGlu 900
 2651 AAGAAACAAAGGAAACCTCAGTCGAGTAGATGATGATGATGATGAT 2700
 901 GluValSerGluAspGlyValSerSerThrValAspGlnLeuSerSp 917
 2701 GAAAGTTTGGAGATGAGATGCTCGAGTCTGTGAGCACTTGTGATCA 2750
 917 eHisIleGluProGlyThrAspAspSerGlnHisSerLysCysAspVal 934
 2751 TCATATAGAGCTCGAGACCATGATTCACGACACTTAATGTGATGAT 2800
 934 sPlySerValGlnProGluProPhePheHisLysValValHisSerG 950
 2801 ATAAAGTCTGTGACCGAGACCATTTTCCATAGGTGTTTCATCTGAA 2850
 951 HisLeuAsnLeuValProGlnValGlnSerValGlnCysSerProGlu 967
 2851 CACTTGAAGCTTATGCTCCCTCAAGTCAATCAAGTCAAGTCAAGTCA 2900
 967 uSerPheAlaPheArgSerHisSerHisLeuProProlLysAsnL 984
 2901 ATCTTGTGATTTGATCTCACTGATTCATTCACCAAAATAATAACA 2950
 984 yAsnSerLeuLeuIleGlyLeuSerThrGlyLeuPheAspAlaAsn 1000
 2951 AGAATTCCTGCTGATGATGATGATGATGATGATGATGATGATGAT 3000
 1001 ProlLysMetLeuArgThrCysSerLeuProAspLeuSerLysLeuP 1017
 3001 CCAAGATGTTAAGACATGTTCACTTCAGATCTCTCAAAAGCTGTTCA 3050
 1017 gThrLeuMetAspValProThrValGlyAspValArgLlnAspAsnLe 1034
 3051 AACCTTATGATGATGTTCCACGCTAGAGATGTTGTCAGCAATCTTG 3100
 1034 lLysAspGluThrLysAspGluAsnIleLysGluIleProSerAspSer 1050
 3101 AAATGATGAAATTAAGATGAAATTAAGATGAAAGGACCTTCTGATCT 3150

1051 GluAspIleValPheGluGluThrThrAspThrAspLeuGlnLysLeu 1067
 3151 GAGACATGTTGTTTGAAGAACTGACACAGATTATCAAGAGCTGACAG 3200
 1067 aSerMetGluGlnLeuLeuArgLlnProGlyGluGluThrThrThrThr 1084
 3201 CTCGATGACAGATGATGATGATGATGATGATGATGATGATGATGAT 3250
 1084 lLysGluSerValLeuLysAsnSerAspValGluProThrAlaAsn 1100
 3251 AAGAGAGTCACTGTTGAGACAGATGATGATGATGATGATGATGATG 3300
 1101 ThrAspValAlaAspGluAspAspAsnProSerSerGluSerAlaLeu 1117
 3301 ACAGATGTCGAGATGAGATGACAAATCCAGATGAGAAAGTCCCTGAA 3350
 1117 nGluGluThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1134
 3351 CGAAGATGCGCATCAGATTAACATGATGATGATGATGATGATGATG 3400
 1134 lLysAspSerValPheAsnHisLeuGlnLysLeuArgLeuHisLeu 1150
 3401 AATGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 3450
 1151 GlnGluMetGlyPheGluLysPhePheGluValTyrGluLysIleLys 1167
 3451 CAGCAATGCGCTTTGAAATTTCTTTGAGTTTATGAGAAATAAGGC 3500
 1167 aLeuHisGluAspGluAspGluAsnIleGluIleCysSerLysIleVal 1184
 3501 TATTCATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3550
 1184 lAsnIleLeuGlnLysnLysGlnHisLeuThrAlaLysIleLeuHis 1200
 3551 AAATATTTTGGAGATGACATGATGATGATGATGATGATGATGATGAT 3600
 1201 LeuValMetAlaAspGlyAlaTyrGlnLysAspAspVal 1214
 3601 TTAGTATGTCAGATGAGACCTCAGAGAGATTAATGATGA 3642

seq_name: gb_pat:AX224518

seq_documentation_block:

LOCUS AX224518 5426 bp DNA linear PAT 10-SEP-2001
 DEFINITION Sequence 50 from Patent WO0161016.
 ACCESSION AX224518
 VERSION AX224518.1 GI:15554730
 KEYWORDS

SOURCE

ORGANISM

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

1 (bases 1 to 5426)

AUTHORS

Walke,D.W., Hu,X., Nepomichy,B., Turner,C.A. and Zambrowicz,B.

TITLE

Human kinases and polynucleotides encoding the same

JOURNAL

Patent: WO 0161016-A 50 23-AUG-2001;
 Lexicon Genetics Incorporated (US)

FEATURES

Location/Qualifiers
 1..5426
 /organism="Homo sapiens"

BASE COUNT

1852 a 927 c 1163 g 1484 t

ORIGIN

alignment_scores:

Quality: 6243.00 Length: 1214
 Ratio: 5.143 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-783-320-4 x AX224518

Align seg 1/1 to: AX224518 from: 1 to: 5426

```

1 MetGluLysTyrValArgLeuGlnLysIleGlyLysSerPheGlyLys 17
|||||
534 ATGGAGAGTATGTTAGACTACAGAAAGATTGGAGAAAGTTTCATTTGGAAA 583
17 sAlaIleuValLysSerThrGluAspGlyArgGlnTyrValIleLysG 34
|||||
584 AGCCATTCTTGTAAATCTACAGAAAGTGGAGACATGATGTTATCAAG 633
34 LuIleAsnIleSerArgMetSerSerLysGluArgGluSerArgArg 50
|||||
634 AAATTAAATCTCAAGATGTCCAGTAAAGAAAGAAAGAAATCAAGAGA 683
51 GluValAlaValIleuAlaAsnMetLysHisProAsnIleValGlnTyr 67
|||||
684 GAAGTTGCAATTTGGCAAAACATGACATCCAAATTTGTCCAGTATAG 733
67 gLysSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
|||||
734 AGAATCATTTGAGAAAGTGGCTCTCTACATGATGATGATGATGATG 783
84 LuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
|||||
784 AGGGAGGGGATCTGTTTAAAGCAATTAATGCTCGAAAGCGCTTTGTTT 833
101 GlnGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeu 117
|||||
834 CAAGAGGATCAGATTTTGGACTGCTTGTACAGATATGTTGGCCCTGAA 883
117 sHisValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsn 134
|||||
884 ACATACATGATGATGAAAAATTTCTTCATCGACATTAATCTCAGACA 933
134 LePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIleAla 150
|||||
934 TATTTTAACTAAAGATGCAACAGTGAACACTTGGAGATTTTGGAAATG 983
151 ArgValIleuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThr 167
|||||
984 AGAGTTCTTAATAGTACCTGAGACTGCTGCAACTTGCATAGAGGACCC 1033
167 cTyrtTyrLeuSerProGluIleCysGluAsnLysProTyrAsnAsnLys 184
|||||
1034 ATACTACTTGTACCTGAAATCTGGAACAACAACCTTACAAATTAATAA 1083
184 eAspIleThrAlaLeuGlnCysValLeuTyrGluLeuCysThrLeuLys 200
|||||
1084 GTGACATTTGGCTCTGGGTGTGCTCTTATGAGCTGTGTACACTTAAA 1133
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValLeuLysIleLeu 217
|||||
1134 CATGCTTTTGAAGCTGCGAGTATGAAAAACCTGTACTGAAAGATAATC 1183
217 rGlySerPheProProValSerLeuHisTyrSerTyrAspLeuArgSer 234
|||||
1184 TGGATCTTTTCCACCTGTGCTTGTGCAATATCTCTATGATCTCCGAGT 1233
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
|||||
1234 TGGTGTCTCAGTATTTAAAGAAATCTCGAGGATAGACATCAGTCAAC 1283
251 SerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPheLeu 267
|||||
1284 TCCATATTGGAGAAAGGTTTATAGCCAAACGCAATGAAAAATTTCTCTC 1333
267 rProGlnLeuIleAlaGlnLysPheCysLeuLysThrPheSerLysPhe 284
|||||
1334 TCCCTAGCTTATTTGCAAGAAATTTGTCTAAAAACATTTTCCAGATTG 1383
284 LysGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
|||||

```

```

1384 GATCAGACCTATACCACTAAAAAGCCAGCTTCAGGACAAAACCTCGATT 1433
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGlyI 317
|||||
1434 TCTGTATTGCTGCTCGAGAAATTTACAAAGCCGCGCTTAATTAATGGAAT 1483
317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGluLysLys 334
|||||
1484 ACCTTTACATATATAGAAATATGAGATTAATAAATTAACAGAAAAAGAAC 1533
334 rOleuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
|||||
1534 CACTGCAAAAAACATTAACAGCCCATCAAACTCCAGAGAAAGAGATGAAT 1583
351 ThrGluGluGluArgLysIleSerGluGluAlaAlaArgLysArgArg 367
|||||
1584 ACTGAGAAAGAAAGAGAAATATCTGAGAAAGCAAGCAAGAAAGAAAG 1633
367 gLeuGluPheIleGluLysGluLysLysGlnLysAspGlnIleIleSer 384
|||||
1634 GCTGCAATTTATGAAAGAAAGAAAGCAAAAGCAATGATGATTAATAGTT 1683
384 euMetLysAlaGluGlnMetLysArgGlnGluLysGluArgLeuGlu 400
|||||
1684 TAAATGAAGGCTGAACAATGAAAAAGCAAGAAAGAAAGGTTGGAAGA 1733
401 IleAsnArgAlaArgGluGlnGlyTyrPArgAsnValLeuSerAlaGly 417
|||||
1734 ATAAATAGGGCCAGGGCAACAGATGAGAAATGTCTTAAGTGTGGTGG 1783
417 ySerGlyGluValLysAlaProPheLeuGlySerGlyGlyThrIleAla 434
|||||
1784 AAGGTGGAAGTAAAGGCTCTTCTGAGGAGTGGAGGAGCTATATAGCTC 1833
434 rOserSerPheSerSerArgGlyGlnTyrGlnHisTyrHisAlaIlePhe 450
|||||
1834 CATCATCTTTTCTTCGAGAGACATATGAACATTAACCATCCATTTT 1883
451 AspGlnMetGlnGlnGlnArgAlaGluAspAsnGluAlaLysTyrLys 467
|||||
1884 GACCAAAATGCAAGCAACAAGACAGAAATATGAAGCTTAATGGAANA 1933
467 gGluIleTyrGlyArgGlyLeuProGluArgGlnLysGlyGlnLeuAla 484
|||||
1934 AGAATATATGTGAGAGTCTCCAGAAAGCAAAAGGCGAGCTAGCTG 1983
484 aGluArgAlaLysGlnValGluGluPheLeuGlnArgLysArgGluAla 500
|||||
1984 TAGAAAGAGCTAAACAAGTGAAGAGTCTCTCAGCGAAACGGAGACT 2033
501 MetGlnAsnLysAlaArgAlaGluGlnHisMetValTyrLeuAlaArg 517
|||||
2034 ATCGAAGATTAAGCTCGAGCCGAGAGACATATGTTTATCTGCAAGACT 2083
517 uArgGlnIleArgLeuGlnAsnPheAsnGluArgGlnGlnIleLysAla 534
|||||
2084 GAGCAAAATGAAGCTACAGAAATTTCAATGAGCGCCAAAGATTAAGCCA 2133
534 yLysLeuArgGlyGluLysLysGluAlaAsnHisSerGlnLysIleGln 550
|||||
2134 AACTTCTGTGTGAAGAAAGAAAGCTATATCATTTGCAAGACAAAGAGA 2183
551 SerGluGluAlaAspMetArgArgLysLysIleGlySerLeuLysAla 567
|||||
2184 AGTAAGAGGCTGACATGAGGCGCAAAATGCAATTCCTGAAGGCCCA 2233
567 sAlaAsnAlaArgAlaAlaValLeuLysGluGlnLeuGluArgLysArg 584
|||||
2234 TGCAAAATGCAAGCTGCTGCTGTACTAAAGAACACATGAAACCAAGAGA 2283
584 yGluAlaTyrGluArgGluLysLysValTyrPoliGlnLysLeuValAla 600
|||||
2284 AGGAGCTTATGAGAGAGAAAAAAGTGTGGAAGAGCATTTGTGTGCT 2333

```

```

601  LysGlyValIysSerSerAspValSerProProlLeuGlnHisGluThr 617
|||||
2334  AAGAGATTAAAGATTCTGATGTTCTCCACTTTGGACAGCATGAAC 2383
|||||
617  rGlyGlySerProSerLysGlnMetArgSerValIleSerValThrS 634
|||||
2384  AGGGGCTCTCCATCAACAGCAACAGATCTGTATTTCTGTAACTT 2433
|||||
634  eAlaLeuLysGluValIglyValAspSerSerLeuThrAspThrArg 650
|||||
2434  CAGCTTGAAGAAGATTGGCGTGCACAGATTAACTGATACCGGGAA 2483
|||||
651  ThrSerGlnMetGlnThrAsnAsnAlaIleSerSerLysArgG1 667
|||||
2484  ACTTCAGAGAGATGCAAAAGACCAACAAATGCTATTCAGTAAACG 2533
|||||
667  uIleLeuArgArgLeuAsnGlnAsnLeuLysAlaGlnGluAspGly 684
|||||
2534  AATCTTCGACAGTAAATGAAATCTTAAAGCTCAAGAGATGAAAG 2583
|||||
684  LysMetGlnAsnLeuSerAspThrPheGlnIleAsnValHisGlu 700
|||||
2584  GAATGCAAAATCTCTGATACTTTTGAGATTAATGTCATGAAATGC 2633
|||||
701  LysGluHisGlyLysGlyLysSerValSerSerAspArgLysTrpG1 717
|||||
2634  AAGAGCATGAAAAAGAAATCAGTTTCATCTGATCCCAAGAGTGG 2683
|||||
717  uAlaGlyGlyGlnLeuValIleProlLeuAspGluLeuThrLeu 734
|||||
2684  GCGAGAGGTCACCTTGATCTCTCGAGTGGATTACACATGATGACT 2733
|||||
734  erPheSerThrThrGluArgHisThrValIglyValIleLysLeu 750
|||||
2734  CCTTCTCAACAACGAAAGACATACAGTGGAGAGCTTATTAATTA 2783
|||||
751  ProAsnGlySerProArgArgAlaTrpGlyLysSerProThrAsp 767
|||||
2784  CCTAATGATCTCCAGAGAGCCTGGGGAAAAAGTCCGACAAATCT 2833
|||||
767  IleuLysIleLeuGlyGlnAlaGluLeuGlnLeuGlnThrGluLe 784
|||||
2834  TCTAAGATACTTGGAGAGCTGAACACTTCAACAGAACTATTTAG 2883
|||||
784  LysAsnThrThrIleArgSerGluIleSerProGlnGlyLysTyr 800
|||||
2884  AAAATACAACTATTGAGAGATTTCTCCGGAAGGGAAAAAGTACAA 2933
|||||
801  ProlLeuIleThrGlyLysLysValGlnCysIleSerHisGluLe 817
|||||
2934  CCTTAAATTACTGAGAAAAAAGTAAATGATTTTCACATGAAATTA 2983
|||||
817  nProSerAlaIleValAspSerProValGluThrLysSerProGlu 834
|||||
2984  CCCATCAGCTATTGTTGATTCCTCGTTGAGACAAAAGTCCCATG 3033
|||||
834  erGlnAlaSerProGlnMetSerLeuLysLeuGlnGlyAsnLeuG 850
|||||
3034  GTGAGGCATCTCCACAGATGTCATTGAAATGGAAGAAATTAAGA 3083
|||||
851  ProAspAspLeuGlnThrGluIleLeuGlnLysProSerGlyThr 867
|||||
3084  CCTGATGATTGGAAAAAGAAATTTCTACAGAGCCAAAGTGAACA 3133
|||||
867  sAspGluSerLeuProCysThrIleThrAspValTrpIleSerGlu 884
|||||
3134  AGATGAGAGCTTGCATGCATATTACTGATGTGGATTTAGTAGAAA 3183
|||||
884  LysGluThrLysGluThrGlnSerAlaAspArgIleThrIleGln 900
|||||
3184  AAGAAACAAAAGAAACTCAGTCGCGAGATAGATCACCATTCAAGA 3233
|||||
901  GluValSerGlnAspGlyValSerSerThrValAspGlnLeuSer 917
|||||
3234  CAGCTTCTGAGAGATGAGCTCGACTACTGTGACCAACTAGTACAT 3283
|||||
917  eHisIleGlnProlGlyThrAsnAspSerGlnHisSerLysCysAs 934
|||||
3284  TCATTAAGCCCTGAGCAACCAATGATTTCTCAGCAGCTTAATGAT 3333
|||||
934  sPylsSerValGlnProGluProPhePheHisLysValValHis 950
|||||
3334  ATAAGCTGTGCAACCGGAACCAATTTTCCATTAAGTGGTTCATT 3383
|||||
951  HisLeuAsnLeuValProGlnValGlnSerValGlnCysSerPro 967
|||||
3384  CACTGAACTTAGTCCCTCAAGTTCAATCAGTTCACTGTTCAACG 3433
|||||
967  uSerPheAlaPheArgSerHisSerHisLeuProPylsAsnLys 984
|||||
3434  ATCCTTTCATTTCCATCTCAGTCTCAGCTTTCACCAAAAATTA 3483
|||||
984  yAsnSerLeuLeuIleGlyLeuSerThrGlyLeuPheAspAla 1000
|||||
3484  AGAATTCCTTCTGATTGGACTTTCACCTGCTGTTGATGCAACAC 3533
|||||
1001  ProlLysMetLeuArgThrCysSerLeuProAspLeuSerLysLe 1017
|||||
3534  CCAAGATGTTAAGGACATGTTCACTTCAGATCTCCAAAGCTGTC 3583
|||||
1017  gThrLeuMetAspValProThrValIglyAspValArgGlnAsp 1034
|||||
3584  AACCTTATGAGATGTTCCACCGTAGAGATGTTGTCAGAACAACT 3633
|||||
1034  LuIleAspGluIleLysAspGlnAsnIleLysGlnGlyProSer 1050
|||||
3634  AATATGATGAATTTAAGATGAAAAACATTAAGAGAGACCTTCTG 3683
|||||
1051  GluAspIleValPheGlnGluThrAspThrAspLeuGlnIleu 1067
|||||
3684  GAAGACATTTGTTTGAAGAAATGACACAGATTTTACAGAGCTCAG 3733
|||||
1067  aserMetGlnLeuLeuAsnArgGlnGlnProGlnGlyGlySer 1084
|||||
3734  CTCGATGGAACAGTTACTTAGGGAACAACTGGTGAAGATACAG 3783
|||||
1084  LuGlnGluSerValLeuLysAsnSerAspValGluProThrAla 1100
|||||
3784  AAGAGAGTCAGCTTGAAGAACAGTGTGAGGCCAACTGCAAATG 3833
|||||
1101  ThrAspValAlaAspGluAspAsnProSerSerGluSerAlaLe 1117
|||||
3834  ACAGATGTGCGAGATGAAAGATGACAACTCCAGTGTGAAGTGC 3883
|||||
1117  nGluGluTrpHisSerAspAsnSerAspGlyGluIleAlaSer 1134
|||||
3884  CGAAGAAATGCACTCAGATMACAGATGAGAAATTTGCTAGTAAG 3933
|||||
1134  LuCysAspSerValPheAsnHisLeuGlnGluLeuArgLeuHis 1150
|||||
3934  AATGCGATAGTCTCTTAACCACTTAAAGAACTAGACTCACTGAG 3983
|||||
1151  GlnGlnMetGlyPheGlnLysPhePheGlnValIleGlyLysI 1167
|||||
3984  CAGAAATGGCTTTGAAAATTTCTTGAGGTTTATGAGAAATAAG 4033
|||||
1167  alIleHisGluAspGluAspGluAsnIleGluIleLysSerLys 1184
|||||
4034  TATTCATGAAAGATGAAAGTAAATTTGAAATTTGTTCCAAATA 4083
|||||
1184  LysAsnIleLeuGlnGlyAsnGlnHisGlnHisLeuThrAlaL 1200
|||||
4084  AAAATATTTTGGGAAGAACATCAGACATCTTTATATGCCAAAG 4133
|||||
1201  LeuValMetAlaAspGlyAlaIleGlnGlnLysAsnAspGlu 1214

```

|||||
4134 TTAGTCATGCGACGACCTACCAAGAGTAATGATGAA 4175

seq_name: gb_AB067488

seq_documentation_block:

LOCUS AB067488 5497 bp mRNA linear PRI 15-SEP-2001
DEFINITION Homo sapiens mRNA for KIAA1901 protein, partial cds.
ACCESSION AB067488
VERSION AB067488.1 GI:15620860

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

BASE COUNT

ORIGIN

alignment_scores:
Quality: 6201.00 Length: 1258
Ratio: 5.112 Gaps: 1
Percent Similarity: 96.423 Percent Identity: 96.343
alignment_block:
US-09-783-320-4 x AB067488
Align seg 1/1 to: AB067488 from: 1 to: 5497

1 MetGluLysTyrValArgLeuGlnLysIleGlyGlnGlySerPheGly 17
|||||
493 ATGAGAGAGATGTGTGACTACAGAGATTGGAGAGGTTCTTGGANA 542
17 SalAlIleuValLysSerThrGluAspGlyArgGlnTyrValIleLysG 34
543 AGCATTCTGTTAAATCTACAGAGATGGCAGACAGTATGTATCAAG 592
34 LuIleAsnIleSerArgMetSerSerLysGluArgGlnGluSerArg 50
593 AAATTACATCTCAGATGTCAGATGCAAGAGAGAGAGAGAGAGAGAG 642
51 GluValAlaValIleuAlaAsnMetLysHisProAsnIleValGlnTyr 67
643 GAATGTCAGTATGGCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAG 692
67 GlnLysPheGlnGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
693 AGAATCATTTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 742
84 LucGlyAspLeuPheLysArgIleAsnAlaGlnGlyValIleuPhe 100
743 AGGAGAGGAGTCTGTTTACGAGATATGCTCAGAGAGAGAGAGAGAGAG 792
101 GlnGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeu 117
793 CAAAGAGATGAGATTTGGAGCTGTTTGAAGATATGTTTGGCCCTGAA 842
117 ShIleValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsn 134
843 ACATGTCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 892
134 LePheLeuThrLysAspGlyThrValGlnIleuGlyAspPheGlyIleAla 150
893 TATTTTAACTAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 942
151 ArgValLeuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThrPr 167
943 AGACTTCTTAAATGATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 992
167 CytTyrLeuSerProGlnIleCysGluAsnLysProTyrAsnAsnLys 184
993 ATCTACTCTGACCTGAATCTGTGAACCAACCTTCAATTAATTAATA 1042
184 erAspIleThrPalaleuGlyCysValIleuTyrGluLeuCysThrLeuLys 200
1043 GTGACATTTGGCTCTGGGCTGTCTTATGAGCTGTGTAACCTTAAA 1092
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValIleLysIleIleSe 217
1093 CATCTTTTGAACCTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1142
217 rGlySerPheProValSerLeuHisTyrSerTyrAspLeuAlaArgSerL 234
1143 TGGATCTTTTCAACCTGCTTTCATTAATCTATGATGATCCCACTT 1192
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
1193 TGGGTCTCACTTATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1242
251 SerIleuGlnLysGlyPheIleAlaLysArgIleGlnLysPheLeuSe 267


```

|||||
1243 TCCATATTGGAGAAAGGTTTATAGCCAAACGATTTGAAAGTTCTCTC 1292
267 rPrGlnLeu1Lea1glu1uPheCysLeu1yThrPheSer1yPheG 284
1293 TCCACACCTATTGACAGAAATTTTCTAAAACATTTTCGAAGTTG 1342
284 1ySerGlnPro1LePro1a1ySarPro1a1SerG1y1nAsnSer1le 300
1343 GATCACAGCCTATACCGTAATAAGACCGCTTCAGACAAAACTCGATT 1392
301 SerValMetPro1a1Gln1y1Ser1yPhePro1a1a1a1y1y1y1 317
1393 TCTGTTATGCCCTGCTCGAATAATTACAAAGCCGCCGCTAAATATGAA 1442
317 eProLeu1a1a1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 334
1443 ACCTTACCATTTAGAAATATGAGATATAAAATTTACAGCAAAAGAAC 1492
384 rOleuGln1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 350
1493 CACTGCAAAAACATTAACAGCCATCACTCAGAGAAAGAGTGAAT 1542
351 ThrG1y1glu1u1a1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 367
1543 ACTGAGAAAGAAAGAGAAATATCTGAGCAAGCAGCAAGAAAGAAAG 1592
367 rLeuGln1uPhe1LeG1u1y1y1y1y1y1y1y1y1y1y1y1y1y1 384
1593 GCTGGAATTTATGAAAAAGAAAAAGAAAGAAAGATTCAGATTATTA 1642
384 eMe1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 400
1643 TAAATGAAGCTGAACCAATGAAGCAAGCAAGAAAGAGTTGGAAAG 1692
401 1Lea1nA1a1a1a1a1a1y1y1y1y1y1y1y1y1y1y1y1y1y1 417
1693 ATAAATATAGGCGAGGAAACAGATGAGAAATGCTGTAAGTCTG 1742
417 ySerG1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 434
1743 AAGTGTGAAGTAAAGGCTCTTTCTGGCAGTGGAGGACTATATAG 1792
434 rOserSerPheSerSerPheSerPheSerPheSerPheSerPheSer 450
1793 CATCATCTTTTCTTCTGAGACAGATATGACATTTACCATGCAATTT 1842
451 AspGlnMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 467
1843 GACCAAAATGACGACAAACAGAGCAAGATATGAGCTTAATGCAAA 1892
467 gGlu1Le1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y 476
1893 AGAAATATATGCTGAGGCTTCCAGAAAGAGAAATTCGCTGAGATT 1942
476 ..... 476
1943 GTCAGAGATTTCCTTATGCGCTGACAGTATCACCATTTTCCGATCT 1992
476 ..... 476
1993 GATGATATTAGAAAACTTTGAAAAGATTGAAGCGGTGCTAAACAA 2042
477 ..... ArgGln1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 490
2043 CATATGCAAAAGGCAAAAGGCGACCTAGCTGTAGAAAAGACCTTA 2092
490 a1G1u1uPhe1uGln1a1y1y1y1y1y1y1y1y1y1y1y1y1y1 506
2093 TAAAGAGTTCTCGACGCAAAAGGGAAGCTATGCAAGATTAAGCT 2142
507 A1G1u1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 523
|||||
2143 GCCAAGACATATGCTTATCTGCAACAGCTGAGCAAAATTAAGACTCA 2192
523 naSPheAsnG1u1a1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 540
2193 GAAATTCAGACGCGCCACACAGATTAAGCCAAATCTGCTGATAAG 2242
540 ySG1u1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1 556
2243 AAGAGCTTAATCATCTGAAAGCAAGAAAGAAAGAAAGAGCTTAC 2292
557 ArgArG1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 573
2293 AGGCGCAAAAAATCGAATCACTGAAGCCCAATGCAAAATGACGCT 2342
573 aValLeu1ySG1u1nLeuG1u1a1y1y1y1y1y1y1y1y1y1y1y 590
2343 TGTACTTAAGAACACACTAGAACGAAAGAAAGAGGCTTATGAGAG 2392
590 1u1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 606
2393 AAAAAAAGTGTGGAGAGACATTTGGCTTAAAGAGTAAAGATTCT 2442
607 AspVal1SerPro1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 623
2443 GATGTTTCTCACCTTTGGGACAGCATGAAGAGTGGCTTCATCA 2492
623 sG1nGlnMetArG1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 640
2493 GCAACAGATGAGATCTGTTATTTCTTAATCTCAGCTTGAAGAA 2542
640 1yVal1Asp1Ser1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 656
2543 GCGTGACAG1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 2592
657 1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 673
2593 AAGACCAACAAATGCTATTTCAAGTACGCAAGAAATCTTGATTA 2642
673 nG1u1a1nLeu1y1a1a1y1y1y1y1y1y1y1y1y1y1y1y1y1 690
2643 TGAATCTTAAAGCTCAAGAAAGATGAAGAAAGAAACCAATCT 2692
690 sP1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y 706
2693 ATACTTTGAGATTAATGCTCATGAGATGCAAGAGCAATGAAG 2742
707 1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 723
2743 AAATCAGTTTCATCTGATCCGACAGAGTGGAGGACGATCACT 2792
723 111ePro1eua1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 740
2793 GATTCCTCGATGATTAACACTGATACATCTCTCTAACAAC 2842
740 rG1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 756
2843 GACATACAGTGGAGAGATTAATTAATTAAGTCTTAATGATCC 2892
757 Arg1a1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 773
2893 AGAGCCTGGGGAAGTCCGACAGATTTCTGTTAAAGATCTGAG 2942
773 u1a1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y 790
2943 AGCTGAACCTACACTTCAACAGACAGATTAATTAATTAATTA 2992
790 e1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 806
2993 GTGAGATTCTCCGAAAGGGAAGTCAAAACCTTATTAATTAAT 3042
807 1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 823
3043 AAAAAATACATATGATTTACATGAATAATAAACCATCATCATTA 3092

```

833 pSerProValGluThrLysSerProGluPheSerGluAlaSerProGluM 840
 3093 TTCTCCTGTTGACAAAAGTCCCGAGTTCAGTGAAGCATCTCCACAGA 3142
 840 eSerLeuLysLeuGluGlyAsnLeuGluLupProAspPheLeuGluThr 856
 3143 TGTCTATGAACGTGAAGAAATTTCAGAACCTGATGATTGGAAACA 3192
 857 GluLeuLeuGluLupProSerGlyThrAsnLysAspGluSerLeuProcy 873
 3193 GAAATCTACAGAGCCAGTGAACAAACAAAGATGAGCTTCCCATG 3242
 873 sThrLeuThrAspValThrPheSerGluGluLysGluThrLysGluThrG 890
 3243 CACTTTCTGATGTGTGATTAGTGAAGAAAGAAAGAAAGAAAGATC 3292
 890 lSerAlaAspArgIleThrIleGluGluAsnGluValSerGluAspGly 906
 3293 AGTCGGGATAGATGATCACCATTGAGAAATGAAGTTCTGAAGATGCA 3342
 907 ValSerSerThrValAspGluLeuSerAspIleHisIleGluProGlyTh 923
 3343 GTCTCGACTACTGTGACCAACTAGTACATTCATATGAGCCCTGGAAC 3392
 923 rAsnAspSerGluHisSerLysCysAspValAspLysSerValGluProG 940
 3393 CAATGATTTCTCAGCAGCTCTAATGTGATGATGATGATGATGATGAT 3442
 940 lUpProPhePheHisLysValValHisSerGluHisLeuAsnLeuValPro 956
 3443 AACCATTTTCCATAGGTGTGTTCTGACACCTTGAACTTGAATGCCCT 3492
 957 GluValGluSerValGluCysSerProGluGluSerPheAlaPheArgS 973
 3493 CAGGTTCATCAGTTCACTGTTCCACGAAAGAAATCCTTTCATTTGATC 3542
 973 rHisSerHisLeuProPheLysAsnLysAsnLysAsnSerLeuLeuIleG 990
 3543 TCACCTGGATTTCACCAAAAATTAACAAAGAAATCTTCTCTCAATG 3592
 990 lLysSerThrGlyLeuPheAspAlaAsnProLysMetLeuArgThr 1006
 3593 GACTTTCAACTGCTCTTTGATGCAAAACCAACCAAGATTTAAGACA 3642
 1007 CysSerLeuProAspLeuSerLysLeuPheArgThrLeuMetAspValPr 1023
 3643 TGTTCACCTCCAGATCTCTCAAAAGCTGTGCAACCTTATGATGATCTC 3692
 1023 OThrValGlyAspValArgGlnAspAsnLeuGluIleAspGluIleLysA 1040
 3693 CACCTAGAGATGTGCTCAAGACATCTTGAATAGATGAATGGAAG 3742
 1040 sPgluAsnIleLysGluGlyProSerAspSerGluAspIleValIlePheGlu 1056
 3743 ATGAAACATTTAAAGAGAGACCTTCTGATCTGAAGCATGTGTTTGA 3792
 1057 GluThrAspThrAspLeuGluGluLeuGluAlaSerMetGluLeuLeu 1073
 3793 GAACTGTACAGATTTACAGAGCTCAGAGCCCTCGATGGAACACATTTAC 3842
 1073 uArgGluGluProGlyGluGluThrSerGluGluGluGluSerValLeu 1090
 3843 TAGGGAACACCTGTGAGAAATACAGTGAAGAAAGAGTCTTGA 3892
 1090 yAsnSerAspValGluProThrAlaAsnGlyThrAspValAlaAspGlu 1106
 3893 AGAACAGATGTGAGCAACTCAATGAGACAGATGTGCGACAGATGA 3942
 1107 AspAspAsnProSerSerGluSerAlaLeuAsnGluLupRphHisSerS 1123
 3943 GATGACATCCACAGTGAAGTGCCTGACAGAAAGATGACATCAGA 3992

1123 pAsnSerAspGluGluIleAlaSerGluCysGluCysAspSerValPhe 1140
 3993 TAACAGTATGAGTGAATTCAGTGAATGATGATGATGATGATGATGAT 4042
 1140 snHisLeuGluGluLeuArgLeuHisLeuGluGluGluMetGlyPheGlu 1156
 4043 ACCATTTAGAGGAATGAGACTTCATCTGACAGAGAAATGAGCTTTGA 4092
 1157 LysPhePheGluValThrGluLysIleLysAlaIleHisGluAspGluAs 1173
 4093 AAATCTTTAGAGTTTATGGAATAATGAAGCTATTCATGAGATGAGA 4142
 1173 pGluAsnIleGluIleCysSerLysIleValGlnAsnIleLeuGluAsnG 1190
 4143 TGAATAATTTGAATAATTTGTCAAAAATAGTCAAAATATTGTGGAAATG 4192
 1190 lHisGluHisLeuGlyThrAlaLysIleLeuHisLeuValMetAlaAspGly 1206
 4193 AACATCGACATCTTATGCCAGATTTCTCATTTAGTCATGCGACATGGA 4242
 1207 AlaTyGluGluAspAsnAspGlu 1214
 4243 GCCTACCAAGAAAGTATATGATGA 4266
 seq_name: gb-pr:HS800531
 seq_documentation_block:
 LOCUS HS800531 5511 bp mRNA linear PRI 18-FEB-2000
 DEFINITION Homo sapiens mRNA; cDNA DKFZ564L2416 (from clone DKFZ564L2416).
 ACCESSION AL050385
 VERSION AL050385.1 GI:4914588
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 5511)
 AUTHORS Koehler,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
 TITLE Direct Submission
 JOURNAL Submitted (15-MAY-1999) MIPs, Am Kiofierspitz 18a, D-82152
 Martinsried, GERMANY
 COMMENT
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by BMFZ (Biomedical Research Center at the Charite,
 Berlin/Germany) within the cDNA sequencing consortium of the German
 Genome Project.
 This clone (DKFZ564L2416) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
 information about the clone and the sequencing project is available
 at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
 Location/Qualifiers
 1..5511
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZ564L2416"
 /clone_1bp="564 (synonym: hfbz2). Vector pAMP1; host
 XL20blue; sites NotI + SalI"
 /dev_stage="etal"
 /rissue_type="brain"
 /map="4"
 join(534..2291,2293..4263)
 /gene="DKFZ564L2416"
 /note="strong similarity to mouse protein kinase nek1
 The frame shift was determined manually"
 /product="hypothetical protein"
 join(534..2291,2293..4263)
 /gene="DKFZ564L2416"
 5476..5481
 polyA_signal
 polyA_site
 BASE COUNT 1883 a 943 c 1190 g 1495 t
 ORIGIN

alignment_scores: Length: 1243
 Quality: 6197.00
 Ratio: 5.109
 Percent Similarity: 97.586 Percent Identity: 97.586

alignment_block:
 US-09-783-320-4 x HSM800531 ..

Align seg 1/1 to: HSM800531 from: 1 to: 5511

```

1 MetGluTyrValAlaArgLeuGlnLysIleGlyGluGlySerPheGly 17
|||||
534 ATGGAGAAAGTATGTAGACACAGAAAGATTGAGAAAGGTTCTTGAAA 583
17 SAIAIleuValLysSerThrGluAspGlyArgGlnTyrValIleLys 34
|||||
584 AGCCATTCTGTATTAATCTACAGAAAGATGCGACAGATGTTATCAAG 633
34 IuIleAsnIleSerArgMetSerSerLysGluArgGluGluSerArg 50
|||||
634 AAATTACATCTCAAGATCTCCAGTAAGAAAGAGAAAGAAATCAAGAG 683
51 GluValAlaValLeuAlaAsnMetLysHisProAsnIleValGlnTyr 67
|||||
684 GAAGTTGCAGATTGGCAAAACATGAAGCATCCAAATATTGTCCAGTAT 733
67 GGIleSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyr 84
|||||
734 AGAATCATTTGAGAAATAATGCTCTCTCATATGAATGATGATTACTGT 783
84 IuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeu 100
|||||
784 AGGGAGGGATCTGTTTAAGCAATTAATGCTCAGAAAGCGGTTTGT 833
101 GInGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeu 117
|||||
834 CAAGGAGATCAAGATTTTGACTGCTGTCTACAGATATGTTGGCCCTGA 883
117 SHISValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsn 134
|||||
884 ACATGTACTGATAGAAAATTTCTTCATCGACATTAATCTCAGAACCA 933
134 IePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIle 150
|||||
934 TATTTTAACTAAGATGAGAACAGTATTAATCTTGAGATTTTGCATTT 983
151 ArgValLeuAsnSerThrValGlnLeuAlaArgThrCysIleGlyThr 167
|||||
984 AGAGTTCTTAATATGACTGTAAGCTGCTGAACTTGCAATGAGGACCC 1033
167 OTYTYrLeuSerProGlnIleCysGluAsnLysProTyrAsnAsnLys 184
|||||
1034 ATACTACTTGTCACCTGAAATCTGAAACAAACCTTACATAATATAAA 1083
184 eArspIleThrAlaLeuGlyCysValLeuTyrGlnLeuCysThrLeu 200
|||||
1084 GTGACATTTGGGCTCGGGGTGTCTCTTATGAGCTGTGATCCCTTAAA 1133
201 HisAlaPheGlnAlaGlySerMetLysAsnLeuValLeuLysIleIle 217
|||||
1134 CATGCTTTGAAGCTGCGAGTATGAAAAAAGCTGACTGAGATAAATATC 1183
217 rGIYserPheProProValSerLeuHisTyrSerTyrAspLeuArgSer 234
|||||
1184 TGGATCTTTCCACCTGCTGCTTTCATATATCTATGATCTCCGAGTT 1233
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
|||||
1234 TGGTGTCTCAAGTATTTAAAGAAATCCAGGAGATAGACATCACTGTAAC 1283
251 SerIleLeuGlnLysGlyPheIleAlaLysArgIleGlyLysPheLeu 267
|||||

```

```

1284 TCCATATTGGAGAAAGTTTATAGCCAAACGATTTGAAAAATTCTCTC 1333
267 rProGlnLeuIleAlaGlnLysPheCysLeuLysThrPheSerLysPhe 284
|||||
1334 TCTCTACCTTATTGCAAGAAATTTGTCTAAAAACATTTTGCAGAGTTG 1383
284 IYserGlnProIleProAlaLysArgProAlaSerGlyGlnAsnSerIle 300
|||||
1384 GATCACAGCCTATACAGCTAAAGACAGCTTACAGACAAACATCCGATT 1433
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGly 317
|||||
1434 TCTGTATTGCTGCTCGAATAATTACAAACCTGCGCTTAATATAGAAAT 1483
317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGlyLysLys 334
|||||
1484 ACCTTACATATATGAATAATGAGATTAATAAATATACAGAAAGAAAC 1533
334 rOLEuGlnLysHisLysGlnAlaHisGlnThrProGlnLysArgValAsn 350
|||||
1534 CACTGCAAAAACATTAACAGGCCCATCAACCTCCAGAGAAAGAGCTGAAT 1583
351 ThrGlyGluGluArgArgLysIleSerGluGluAlaAlaArgLysArg 367
|||||
1584 ACTGGAGAGAAAGAGAGAAATATCTGAGAGAGCGACAGAAAGAGAG 1633
367 GLeuGluPheIleGlyLysGlnLysLysGlnLysAspGlnIleIleSer 384
|||||
1634 GCTGGAATTTATTGAAAAAGAAAGAAACAAAGAGATCAGATTATAGTT 1683
384 eUMetLysAlaGlnLysArgGlnLysArgGlnLysGluArgLeuGluArg 400
|||||
1684 TATATAGCGCTGACAAATGAAAGGCAAGAAAGCAAGAGTTGCAANA 1733
401 ILeaAsnArgAlaArgGlnGlnLysTyrArgAsnValLeuSerAlaLys 417
|||||
1734 ATAAATAGGCGCAGAGAACAGATGGAATAATGTGTAAGTCTGCTGTGG 1783
417 YSerGlyGluValLysAlaProPheLeuGlySerGlyGlyThrIleAla 434
|||||
1784 AAGTGTGAAGTAAGGCTCTCTTCTGCGCAGTGGAGGACATATAGCTC 1833
434 rOserSerPheSerSerArgLysGlnTyrGlnHisTyrHisAlaIlePhe 450
|||||
1834 CATCATCTTTTCTCTCCAGAGACAGTATGAACATTAATGAACTTAATG 1883
451 AspGlnMetGlnGlnArgAlaGlnAspAsnGlnAlaLysThrLysArg 467
|||||
1884 GACCAATGACAGCAACAAAGACAGACATATATGAACTTAATGCAAAAG 1933
467 gGIuIleTyrGlyArgGlyLeuProGluArgGlnLysGlyGlnLeuAla 484
|||||
1934 AGAATATATGTCGAGGCTCTCCAGAAAGCAAAAGGCGAGCTAGCTG 1983
484 aLIGluArgAlaLysGlnValGlnLysPheLeuGlnAlaArgLysArgLys 500
|||||
1984 TAGAAAGAGCTAAACAAAGTAGAAGATTCTCGACGCGAAAGCGGAGCT 2033
501 MeGlnAsnLysAlaArgAlaGlnLysHisMet..... 511
|||||
2034 ATGCAATATTAAGCTGACAGCCAGAGACATATGGAAATCCGCAAAAC 2083
511 ..... 511
2084 GCGAGCTATGTATGAGAGCCAGGCCACCTTTCAAGAGAGAGAGCAAA 2133
512 .....ValTYrLeuAlaArgLeuArgIleArgLeu 522
|||||
2134 GAAACAAAGAGAGAGGTTTATCTGCAAGACTAGGCAATTAAGACTA 2183
523 GInAsnPheAsnGlnLysGlnGlnIleLysAlaLysLeuArgGlyGly 539
|||||
2184 CAGATTTCAATGAGCCCAACAGATTAAAGCAAACTTCTGTTGGAANA 2233

```

539 sLYSGluAAsnHisSerGluGluGluGluGluSerGluGluGluAAspK 556
|||||
2234 GAAGAAGACCTAATCATCTCTGAAGAGCAAGAAGAGAGAGAGCTGAC 2283
556 eTArgArg.LysLysIleGluSerLeuLysAlaHisAlaAsnAlaArgAl 572
|||||
2284 TGAGGCGCAAAAAAATGATCATCTGAAGCCCATGCAATGCACGTGC 2333
572 aAlaValLeuLysGluGluGluGluGluGluGluGluGluGluGluGlu 589
|||||
2334 TCCTCTACTAAAGACACACTAGAACCAAGAAAGAGAGCTTATGACA 2383
589 rGGLuLysValTProGluHisLeuValAlaLysGlyValLysSer 605
|||||
2384 GGAAGAAAGAGTGGAGAGCAATTTGGCTTAAGAGATTAAAGAGT 2433
606 SerAspValSerProProLeuGlyGluHisGluThrGlyLysProSe 622
|||||
2434 TCTGATGTTTCTCCACCTTTGGAGACAGCAAGAAAGAGTGGCTCATC 2483
622 rLYSGluGlnMetArgSerValIleSerValThrSerAlaLeuLysGluY 639
|||||
2484 AAGCAACAGATGAGATCTGTTATTTCTGTAACCTTCAGCTTGAAGAAG 2533
639 aGlyValAspSerSerLeuThrAspThrArgLutThrSerGluGluMet 655
|||||
2534 TTGGCGTGACAGTAGTTTAACTGATACCGGGAACCTTCAGAAAGAGT 2583
656 GlnLysThrAsnAsnAlaIleSerSerLysArgLutIleuArgArgLe 672
|||||
2584 CAAAAGACCAACATGCTATTTCAGTAACGAGAAATACTTGGCAGATT 2633
672 uASnGluAsnLeuLysAlaGlnGluAspGluLysGlyMetGlnAsnLeu 689
|||||
2634 AATGAAATCTTAAAGCTCAAGAGAGTAAAGAAAGAAATGCAAAATCTCT 2683
689 eRAspThrPheGluIleAsnValHisGluAspAlaLysGluHisGluLys 705
|||||
2684 CTGATACCTTTGAGATAAATGTCATGAAAGATGCCAAAGAGCAATGAAAA 2733
706 GlnLysSerValSerSerAspArgLysLysTrpGluAlaGlyGluLe 722
|||||
2734 GAAAAATCAGTTTCATCTGATCCAGAAAGTGGAGGAGGAGGAGTCAACT 2783
722 uValIleProLeuAspGluLeuThrLeuAspThrSerPheSerThrThrG 739
|||||
2784 TGTGATTCCTCGATGAGTTAACACTAGATACATCTCTCTCAACAAC 2833
739 LuArgHisThrValGlyGluValIleLysLeuGluProAsnGlySerPro 755
|||||
2834 AAAAGCATACAGTGGAGAACTATTAAATTAGTCTTAATGATGATCCCA 2883
756 ArgArgAlaTrpGlyLysSerProThrAspSerValLeuLysIleLeuG 772
|||||
2884 AGAAGAGCGCTGGGAAAGTCCGACAGATTCTGTTAAAGATACCTGG 2933
772 yGluAlaGluLeuGlnLeuGlnThrGluLeuGluAsnThrThrIleA 789
|||||
2934 AGAAGCTGACACTACAACTTCAGACAGAACTATTGAAAAATACAACTATTA 2983
789 rGSeGluIleSerProGluGlyGluLysTrpLysProLeuIleThrGly 805
|||||
2984 GAAAGTGAGATTCTCCGAGAGGGAAGATACAAACCTTAATTAAGTGA 3033
806 GlnLysLysValGlnCysIleSerHisGluIleAsnProSerAlaIleVa 822
|||||
3034 GAAAAAGAGTACATGATTTCACATGAAATAAACCATCAGCTATGTT 3083
822 LAspSerProValGluThrLysSerProGluPheSerGluAlaSerProG 839
|||||
3084 TGATTCTCTGTTGAGCAAAAAGTCCGAGTTTCAAGTGAAGCAATCTCCAC 3133

839 lMetSerLeuLysLeuGluGlyAsnLeuGluGluProAspAspLeuGlu 855
|||||
3134 AGATGTCATTGAAAGCTGAAAGAAATTTAGAAAGCAACCTGATGATTGGAA 3183
856 ThrGluIleLeuGlnGluProSerGlyThrAsnLysAspGluSerLeuPr 872
|||||
3184 ACAGAAATTTCTACAGAGCCCAAGTGGACAAACAAAGATGAGAGCTTCC 3233
872 oCysThrIleThrAspValTrpIleSerGluGluLysGluThrLysGluT 889
|||||
3234 ATGCACTATTACTGATGTGTGGATTAGTGAAGAAAAAGAAACAAAGAAA 3283
889 hrcGlnSerAlaAspArgIleThrIleGlnGluAsnGluValSerGluSp 905
|||||
3284 CTCAGTGGCAGATAGGATCCCATTCAGAAAGAAAGTTCCTGAAGAT 3333
906 GlyValSerSerThrValAspGlnLeuSerAspIleHisIleGluProG 922
|||||
3334 GGAATCTCGAGTACTGTGACCAACTTAAGTATGACATTCAATAGAGCCCTG 3383
922 yThrAsnAspSerGlnHisSerLysCysAspValAspLysSerValGlnP 939
|||||
3384 AACCAATGATTCTCAGCACACTTAATGATGATGATGATGATGATGATGAT 3433
939 rGluLupProPhePheHisLysValValHisSerGluHisLeuAsnLeuVal 955
|||||
3434 CGGAACCATTTTCCATTAAGTGTTCATTTCTGAACACTTGAACCTTAGTC 3483
956 ProGlnValGlnSerValGlnCysSerProGluGluSerPheAlaPheAr 972
|||||
3484 CCTCAATTCATCAATCAGTTCAGTGTTCACCAAGAAATCCTTGCATTTCG 3533
972 gSerHisSerHisLeuProProLysAsnLysAsnLysAsnSerLeuLeuI 989
|||||
3534 ATCTCACTCCCATTTACCCCAAAAATTAACAAAGATTTCTCTTGCGA 3583
989 lGlyLeuSerThrGlyLeuPheAspAlaAsnAsnProLysMetLeuArg 1005
|||||
3584 TTGGACCTTTCACTGCTGCTGTTGATGCAAAACCAACCAAGATGTTAAGG 3633
1006 ThrCysSerLeuProAspLeuSerLysLeuPheArgThrIleMetAspVa 1022
|||||
3634 ACATGTTCACTTCACAGATCTCTCAAGCTGTTCAAGACCTTATGAGATT 3683
1022 lProThrValGlyAspValArgGlnAspAsnLeuGluIleAspGluIleL 1039
|||||
3684 TCCACCGCTAGGAGATGTTGTCAGACAAATCTTGAAATAGATGAATTGA 3733
1039 yAspAspGluAsnIleLysGluGlyProSerAspSerGluAspTrIleValPhe 1055
|||||
3734 AAGATGAAACATTAAGAAAGACCTTCGATTCGAAACATTTGTGTTT 3783
1056 GlnGluThrAspThrAspLeuGlnGluLeuGlnAlaSerMetGlnGluLe 1072
|||||
3784 GAAAGAACTACACAGATTTACAAAGCTGCAAGGCTCATGTAACAGATT 3833
1072 uLeuArgGluGlnProGlyGluGluLutLysSerGluGluGluSerValL 1089
|||||
3834 ACTTAGGGAACAACCTGCTGCAAGAAATACAGTGAAGAAAGAGTGAAGTCT 3883
1089 eLysAsnSerAspValGluProThrAlaAsnGlyThrAspValAlaAsp 1105
|||||
3884 TGAAGACACAGTATGATGAGCCCAACTGCAAAATGGGACAGATGTGCAAGAT 3933
1106 GlnAspAspAsnProSerSerGluSerAlaLeuAsnGluGluTrpHisSe 1122
|||||
3934 GAAATGACAAATCCAGTACTGAAAGTCCCTGAAGAAAGAAATGACACTC 3983
1122 rAspAsnSerAspGlyGluIleAlaSerGluCysGluCysAspSerValP 1139
|||||
3984 AGATTAACAGTATGATGTAATTTCTAGTGAATGTGAATGATGATGATGAT 4033
1139 hAsnHisLeuGluGluLeuArgLeuHisLeuGluGlnGluMetGluPhe 1155

```

|||||
4034 TTTAACCTTTTGAAGAACTGAGCTTCATCTGAGCAAGAAATGGCTTT 4083
1156 G1u1ySPhSpheGluValTyrg1u1yS1le1ySAla1le1n1S1gluAspG1 1172
|||||
4084 GAAAAATCTTGTGAGTTATGAGAAATAAAGGCTATTTCATGAGAAGA 4133
1172 uASpG1uAsn1leG1u1leCySer1yS1leValG1nAsn1leuG1yA 1189
|||||
4134 AGATGAAATATTGAAATTTGTTCAAAATAGTCAAAATATTTTGGGAA 4183
1189 snG1u1n1S1n1S1leuTyrg1Ala1yS1leu1n1S1leuValMetAlaSp 1205
|||||
4184 ATGACATCAGCATCTTATGCAAGATCTTCATTACTATGTCAGAT 4233
1206 G1yAlaTyrg1nG1uAspAsnAspG1u 1214
4234 GGAGCTTACCAAGAAATATGATGAA 4260

seq_name: gb_pat:AX224473

seq_documentation_block:
LOCUS AX224473 3024 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 5 from Patent WO0161016.
ACCESSION AX224473
VERSION AX224473.1 GI:15554707
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3024)
Wolke,D.W., Hu,Y., Nepomichy,B., Turner,C.A. and Zambrowicz,B.
Human kinases and polynucleotides encoding the same
JOURNAL Patent: WO 0161016-A 5 23-AUG-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..3024
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1115 a 503 c 681 g 725 t
ORIGIN

alignment_scores:
Quality: 5162.00 Length: 1007
Ratio: 5.126 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-783-320-4 x AX224473 ..

Align seg 1/1 to: AX224473 from: 1 to: 3024

208 Met1yAsn1leuVal1leu1yS1le1leSerG1ySerPhSpheProValSe 224
1 ATGAAAAACCTGCTACTGAAAGTAATATCTGATCTTTTCCACCTGTGTC 50
224 r1eu1n1S1TySerTyfAsp1euArSer1euValSerG1n1eu1eu1e1ySA 241
|||||
51 TTTGCAATATATCTATGATCTCCGACAGTTTGCTGCTCAGTTATTTAAAA 100
241 rGAsnProArGAspArGProSerValAsnSer1le1euG1u1yG1yPhe 257
|||||
101 GAATCTTAGGATAGACATCAGTCAACTCCATTCATTTGAGAAAAGGTTTT 150
258 1le1Ala1ySArG1leG1u1ySPhSphe1euSerProG1n1eu1le1AlaG1uG1 274
|||||
151 ATACCAAAAGCATTTGAAAAGTTTCTCTCTCCTCAGCTTATTTGCAAGA 200
274 uPhCyS1eu1yS1ThrPhSer1ySPhG1ySerG1nPro1leProAla1 291
|||||
201 ATTTGCTATAAAACATTTTGGAGAGTTTGATCACAGCCTATACCAAGCTA 250

```

Wed May 15 14:13:58 2002

us-09-783-320-4.rge

Page 12

591 yslvsvaltrpglugluhislleuvalalalsgylvallysserserasp 607
|||||
1151 AAAAGGTGGGAAGACATTTGGCTAAGAGAGTTAAGAGTTCAT 1200
608 valserproleuglycinhisgluthrclglyserproserlysgl 624
|||||
1201 GTTCTCCACCTTTGGACACATGACAGAGTGGCTCTCCATCAGCA 1250
624 ngimelargservallieservalthrseralaleuylsglvalgl 641
|||||
1251 ACAGATGAGATCTGTTATTTCTGTAACCTGAGCTTGAAGAAAGTGGC 1300
641 alaspserleuthrthrargluthrsergluglumetglulys 657
|||||
1301 TGGACAGATGTTTAACTGATACCGGGAACTTCAGAAAGATGCAGAA 1350
658 thrasnasnalalieserSerlysrargluileuarargleuansgl 674
|||||
1351 ACCAACATGCTATTTCAGTAAAGCAAGAAATCTTCAGATTTAAATGA 1400
674 usasnleuylsalaglugluaspglulysgllymetglinasnleuaserasp 691
|||||
1401 AATCTTAAAGCTCAAGAAATGAAAAAGAAATGCAGAAATCTCTGATGA 1450
691 hrpheglileasnvalhsgluaspalalysgluhisglulysglulys 707
|||||
1451 CTTTGAATTAATTTTCATGACATGCCAAGAGCATGAAAAAGAAAA 1500
708 servalserSerasparglysertrpgluahisglulysglulys 724
|||||
1501 TCGAGTTTCATGATCGCAAGAGTGGAGAGAGAGTCAAGCTGATGAT 1550
724 erproleuaspgluleuthrleuaspthrserpheserthrthgluargh 741
|||||
1551 TCCCTGGAGAGTTAACACATGATACCTCTCTACAACTGAAGAAC 1600
741 lsfthrvalglgluvalilleylsleuglyproasnlyserproargarg 757
|||||
1601 ATACAGTGGAGAAAGTTATTAATTAAGTCTTAATGAGATCTCCAAAGAA 1650
758 alatrpglylsrserprothrAspserValleuylsilleuuglyglua 774
|||||
1651 GCCTGGGGAAAAAGTCCGACAGATTCGTCTTAAAGATACTGGAGAGC 1700
774 agluleugluileuglnthrgluleuyluasnthrthrleargserg 791
|||||
1701 TGAACACACTTCAGACAGAACTATTGAAAAATACAACTATTGAAAGTG 1750
791 luileserprogluglylulysrlysrproleuilethrnglylulys 807
|||||
1751 ACATTTCTCCGAAAGGAGGAAAAAGTCAAAACCTTAATTACTGAGAAAA 1800
808 lysvalglnlyslleserhisglulileasnproseralalevalaspse 824
|||||
1801 AAAGTACAAATGATTTTCATGAATAAACCACATGATATGTTGATTC 1850
824 rprovalgluthrlyserproglupheserglualiaserproglumets 841
|||||
1851 TCCCTGTGAGCAAAAAAGTCCCGAGTTCAGTGAAGCATCTCCACAGATGT 1900
841 erleuylsleugluileuasnleugluugluaspaspaspserleuThglu 857
|||||
1901 CATTTAAACTGGAAAGAAATTTAGAAACCTGATGATTTGAAACAGAA 1950
858 illeleuglnluprosergllythrAsnlysrpcluserleuprocyrst 874
|||||
1951 ATTTCTACAAAGACCAAGTGAACAAACAAAGATGAGAGCTTGGCATGCAC 2000
874 rlllethrAspvaltrpilieserglulysgluthrlysgluthrclns 891
|||||
2001 TATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2050
891 erhlasparglilethrilleuglnuasngluvalisergluaspgllyal 907
|||||

|||||
2051 CGGCAGATAGATCACCATTACAGAAATGAAAGTTCTGTAAGATGAGATC 2100
908 serSerthrvalaspgluleuSeraspilileisilegluproglythrAs 924
|||||
2101 TCGAGTACTGTGACCACTTATGACATTCATATGAGACCTGGAAACCA 2150
924 naspserrghisSerlyscysaspvalaspysSerValglnproglup 941
|||||
2151 TGATTTCTACACTCTTAATGTGATGATGATGATGATGATGATGATGAT 2200
941 rophehisllyslvalhislsergluhislleuasnleuvalproglu 957
|||||
2201 CATTTTCCATTAAGGTGGTTCATTCGAACTGAACTGAACTGATCCCTCA 2250
958 valglinseryalincysSerproglugluserphealapheargserrh 974
|||||
2251 GTTCAATCAGTTCAGTGTTCACCAAGAAATCCTTTCATTCGATTCGA 2300
974 sSerhisleuProprolysrAsnlysrAsnlysrAsnserleuleuilegl 991
|||||
2301 CTCCCATTTACACCAAAAAATAAAAACAAGATTCCTTGATTTGGAC 2350
991 euserthrnglyleupheaspalaspasnprolysmeltheargthrlys 1007
|||||
2351 TTTCACTGGTCTGTTGATGCCAAACACCAAGATGTTAAGGACATGT 2400
1008 serleuProaspserlysrleupheargthrleuMetAspvalproth 1024
|||||
2401 TCACCTTCAGATCTCTTAAGCTGTTCCAGACCCCTTATGATGATGCCAC 2450
1024 rValgllyAspvalarglnaspasnleugluileaspglulileylsasp 1041
|||||
2451 CGTAGAGATGTTCTGTAAGACATCTTGAAGATGATGATGATGATGATG 2500
1041 luasnileylsgluglyproSerAspsergluaspilleyalpheglu 1057
|||||
2501 AAAACATTAAGAAAGACCTTCGATTCGAAACATTCGTTTGAAGAA 2550
1058 thrAspThrAspLeuglnleuglnalasermetglulnleuuar 1074
|||||
2551 ACTGACACAGATTTTACAAAGCTGCAGGCCCTCATGGAACAGTACTAG 2600
1074 ggluglnproglugluuyltrserglugluglugluserValleuLysA 1091
|||||
2601 GCAACAACCTGGTGAAAGATACAGTGAAGAAAGAGTCACTGCTTGAAGA 2650
1091 snseraspvalgluprothrAlasnglythrAspvalAlaspclusp 1107
|||||
2651 ACAGTGAATGTGGAGCCAACTGCAAAATGGGACAGATGTGGCAATGAAT 2700
1108 aspAsnproserSergluseralaleuasnglugluThrhisSerAspAs 1124
|||||
2701 GACATCTCCAGTGTGAAGAGTCCCTGAAACAGAAATGACACTCAATTA 2750
1124 nserAspglgluilealaserglucylsaspserValpheasn 1141
|||||
2751 CAGTGAATGGTGAATTTCTAGTAATGTGAATCCGATGTGCTTTAAC 2800
1141 lslleuglugluuargleuasnleuuglunlmetgllyrphelugly 1157
|||||
2801 ATTTAGAGGAAGTGAAGACTTCATCTGGAGAGGAAATGGCGCTTTGAAA 2850
1158 phephegluvaltyrglylulysilleylAlalehisgluaspclusp 1174
|||||
2851 TTTCTTGGAGTGTATGAGAAATTAAGGCTATTCATGAGATGAGAGATGA 2900
1174 uasnileglulilecysSerlysllevalglnasnilleuuglyAsnnglu 1191
|||||
2901 AAATATTGAATTTGTTCAAAAATAGTCAAAATATTTTGGGAAATGAA 2950
1191 lsglnhisleuTyralalysilleuasnleuvalmetAlaspgllyal 1207
|||||

```

2951 ATCAGACTTTTATGCGCAGATTCTTCATTAGTCATGCGAGATGAGGCC 3000
1208 TGTGInGUAspAsnAspGlu 1214
|||||
3001 TACCAGAAAGATATGATGCAA 3021

seq_name: gb_pat:AX224469

seq documentation block:
LOCUS AX224469 3108 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 1 from Patent WO0161016.
ACCESSION AX224469
VERSION AX224469.1 GI:15554705
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 3108)
AUTHORS Walke,D.W., Hu,Y., Nepomitchy,B., Turner,C.A. and Zambrowicz,B.
TITLE Human kinases and polynucleotides encoding the same
JOURNAL Patent: WO 0161016-A 1 23-Aug-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1. 3108
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1145 a 520 c 708 g 735 t
ORIGIN

alignment_scores:
Quality: 5138.00 Length: 1035
Ratio: 5.102 Gaps: 1
Percent Similarity: 97.295 Percent Identity: 97.295

alignment block:
US-09-783-320-4 x AX224469 ..

Align seg 1/1 to: AX224469 from: 1 to: 3108

208 MetLysAsnLeuValIleuLysIleIleSerGlySerPheProProValSe 224
|||||
1 ATGAAAGAACTGGTACTGAGATATATCTGGATCTTTTCCACCTGTGTC 50

224 rLeuHisTyrSerTyrAspLeuArgSerLeuValSerGlnLeuPheLysA 241
|||||
51 TTTGGATTATTCCTATGATCTCCGAGTTGGTGTCTCACTTATTTAAAA 100

241 rGAsnProArgAspArgProSerValAsnSerIleLeuGlnLysGlyPhe 257
|||||
101 GAAATCTCAGGATAGACCACTGACCACTCCATATTGGAGAAAGGTTT 150

258 IleAlaLysArgIleGlnLysPheLeuSerProGlnLeuIleAgluG 274
|||||
151 ATAGCCAAACGCATTGAAAGATTTCTCTCTCCAGCTTATTCGAGAA 200

274 uPheCysLeuLysThrPheSerLysPheGlySerGlnProIleProAla 291
|||||
201 ATTTGTCTTAAACATATTCGAACTTGGATCAGACCTATATCCACACTA 250

291 ySarArgProAlaSerGlyGlnAsnSerIleSerValMetProAlaGlnLys 307
|||||
251 AAAGACCAGCTTCAGACAAACTCGATTTCTGTTATGCCGCTCAGAAA 300

308 IleThrLysProAlaIleAlaLysTyrGlyIleProLeuAlaTyrLysLys 324
|||||
301 ATTACAAAGCGCTGCCCTAAATATGAAATACCTTTAGCATATTAAGAA 350

324 rGlyAspLysLysLeuHisGlnLysLysProLeuGlnLysHisLysGln 341
|||||
351 TGGAGATTAATAAATTACAGAAAAAAGAAACCACTGCAGAAAAACATAAAC 400

```

```

341 IahISGlnThrProGlnLysArgValAsnThrGlyGlnGluArgArgLys 357
|||||
401 CCCATCAAACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 450
|||||
358 IleSerGlnGlnLysAlaIleArgLysArgArgLeuGlnPheIleGlyLys 374
|||||
451 ATATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 500
|||||
374 uLysLysGlnLysAspGlnIleIleSerLeuMetLysAlaGlnMetL 391
|||||
501 AAAAGAAACAAAGAGATCAGATTATTTATGAAAGCTGACCAATATGA 550
|||||
391 ySarArgGlnLysGluArgLeuGluArgIleAsnArgAlaArgGlnLys 407
|||||
551 AAAGCAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
|||||
408 GLTTPArgAsnValLeuSerAlaGlyIleSerGlyGlnLysAlaArg 424
|||||
601 GGATGAGAGAAATGCTGAAAGCTGCTGGTGAAGTGTGAAGTAAAGCT 650
|||||
424 oPheLeuGlySerGlyGlyThrIleAlaProSerSerPheSerSerArg 441
|||||
651 TTTTCTGGCGAGTGGAGGACATATGCTCATCATCTTTTCTCTCGAG 700
|||||
441 yGlnIleTyrGlnHisTyrHisAlaIlePheAspGlnMetGlnGlnArg 457
|||||
701 GACAGTATGAAACATTTACCATGCTTTTGGACCAATGACAGCAAGAA 750
|||||
458 AlaGlnAspAsnGlnLysAlaLysTyrPheArgGluIleTyrGlyArg 474
|||||
751 GCGAAAGATATATGAAGCTAAATGSAAGAAAGAAATATATGTCGAG 800
|||||
474 uProGlnArgGlnLysGlyGlnLeuAlaValGluArgAlaLysGlnVal 491
|||||
801 TCCAGAAAGCGCAAAAGGCGAGCTAGCTGTGAAAGAGCTAAACAAG 850
|||||
491 IuGlnPheLeuGlnArgLysArgLysAlaMetGlnAsnLysAlaArgAla 507
|||||
851 AAGAGTTCTCGAGCGAAGAGGAGACCTATGCAATTAAGCTCGAGGC 900
|||||
508 GlnGlyHisMet..... 511
|||||
901 GAAGGACATATGGAATCTCTGCAAAACCTGGACGCTATGATGAGGAG 950
|||||
512 .....ValT 513
|||||
951 GCCACCTCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1000
|||||
513 yLeuAlaArgLeuArgGlnIleArgLeuGlnAsnPheAsnGluArgGln 529
|||||
1001 ATCTGCAAGACTGAGAGCAAAATAGACTACAGAAATTCATGAGCGCCA 1050
|||||
530 GlnIleLysAlaLysLeuArgGlyGlnLysLysGlnLysAlaAsnHisSer 546
|||||
1051 CAGATTAAAGCCAAACTCTGCTGTGTAAGAAAGAAAGAAAGCTATCT 1100
|||||
546 uGlyGlnGlnLysSerGlnGluAlaAspMetArgArgLysLysIleGlu 563
|||||
1101 AGGACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1150
|||||
563 eLeuLysAlaHisAlaAsnAlaArgAlaValAlaLysGlnLysGln 579
|||||
1151 CACTGAAGGCCCATGCAATGACAGCTGCTGCTGACTTAAAGAACACTA 1200
|||||
580 GlnArgLysArgLysGlnLysAlaTyrGlnArgGlnLysLysValITP 596
|||||
1201 GAACGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1250
|||||
596 uHisLeuValAlaLysGlyValLysSerSerAspValSerProLeu 613
|||||
1251 GCATTGTGGTCTTAAAGAGATTAAAGGTTCATGATGTTCTCCACCTT 1300
|||||
613 yGlnHisGlnThrGlyGlySerProSerLysGlnGlnMetArgSerVal 629

```

```

|||||
1301 GACACGATGAAACAGTGGCTCTCCATCAAGACAGATGAGATCTGTT 1350
630 IIESeValIThrSerAlaLeuLysGIuValGIValIAspSerSerLeuTh 646
1351 ATTCTGTAGCTTCAGCTTTGAAGAAGTTGGCGTGACAGTGGTTAAC 1400
646 rASpThrTgGIuThrSerGIuLumetGIuLysThrASnAlaIles 663
1401 TGATACCCGGGAACCTTCAGAAAGATGCCAAAGACCAACAATGCTATT 1450
663 eSeSerIArgGIuIleLeuArgArGIeunGIuAsnLeuLysAlaGIu 679
1451 CAAGTAAGCGAANAATCTTCGAGATTAAATGAAATCTTAAAGCTCA 1500
680 GIuAspGIuLysGIuMetGIuAsnLeuSerASpThrPheGIuIleAsnVa 696
1501 GAAGATGAAAAAGAAATGCAGAAATCTCTGATACCTTTGAGATTAATGT 1550
696 IHisGIuAspAlaLysGIuHisGIuLysGIuLysSerValSerSerAspa 713
1551 TCATGAAAGATGCCAAAGAGCATGAAAAAAGAAATCAGTTTCATCTGATC 1600
713 rGLysLysTrpGIuAlaGIuLysGIuLeuValIleProLeuAspGIuLeu 729
1601 GCAAAGATGGGAGGACGAGGATGACCTTGATTCCTCTGATGAGTTA 1650
730 ThrIleuAspThrSerPheSerThrThrGIuArGIHisThrValGIuLys 746
1651 ACAGTAAGTACATCTCTCTACAACTGAAAGACATACACTGGAGAGAGT 1700
746 ILeuLysLeuGIuProAsnGIuSerProArGIAlaTrpGIuLysSerP 763
1701 TATTAAATTAGTCTCTAATGATCTCCAGAAAGCCTGGGGGAAAAAGTC 1750
763 rOTHisASpSerValLeuLysIleLeuGIuLysGIuLeuGIuLeuGIu 779
1751 CGACAGATCTGTCTAAGATCTGAGAGCTGAACTACAACTCAAC 1800
780 ThrGIuLeuLeuGIuAsnThrIleArgSerGIuIleSerProGIuGI 796
1801 ACAGACATATTAGAAATATACAACTATTAGAGATGATTTCTCCCGAAG 1850
796 YGIuLysTrpLysProLeuIleThrGIuLysLysValGIuLysIles 813
1851 GGAAGAAATGAAACCCCTTAATTACTGAGAAAAAAGTACAATGTATTT 1900
813 eRHISGIuIleAsnProSerAlaIleValAspSerProValGIuThrLys 829
1901 CACATGAATATTAACCCATCAGCTATTGTGATTCCTCTGTGAGACAAA 1950
830 SerProGIuLysSerGIuAlaSerProGIuMetSerLeuLysLeuGIu 846
1951 AGTCCCGATTCAAGAGCATCTCCACAGATGTCATTGAAACTGGAGAG 2000
846 YAsnLeuGIuGIuProAspAspLeuGIuThrGIuIleLeuGIuGIuPro 863
2001 AATATTAGAAACCTGATGATTGGAAACAGAAATTCCTACAAAGCCAA 2050
863 eGIuIThrAsnLysAspGIuSerLeuProCysThrIleThrAspValTrp 879
2051 GTGGACAAACAAAGATGAGAGCTTGCATGCTACTATTACTGATGTGG 2100
880 IIESeGIuLysGIuLysGIuThrLysGIuLysGIuInserAlaAspArgIleTh 896
2101 ATTACTGAGGAAAAAGAAACAAAGAAACCTCAGTCGCGACATGAGATC 2150
896 rIleGIuLysGIuValIleSerGIuAspGIuValSerSerThrValAsp 913
2151 CATTCAGGAAATGAAATTTCTGAAGATGAGAGCTCGAGTACGTGTGACC 2200
913 ILeuSerAspIleHisIleGIuProGIuThrAsnAspSerGIuHisSer 929
|||||

```

```

2201 AACTAGTCACATTCATATTAGAGCCCTGAAACCAATGATTCAGACACTCT 2250
930 LysCysAspValAspLysSerValGIuProGIuLysProPheHisLysVa 946
2251 AAATGTGATGATAGATTAAGTCTGTGCACCGAACCAGATTTTTCATTAAG 2300
946 YValHisSerGIuHisLysAsnLeuValIProGIuValInserValGIu 963
2301 GGTTCATTCGTGAACACTTGAACCTAGTACCTCCCTCAAGTTCATACAGT 2350
963 YSerProGIuLysSerPheAlaPheArgSerHisSerHisLeuProPro 979
2351 GTTCACCGAAGAAATCTTTGCATTTCCATCTCCTCCATTTACACCA 2400
980 LysAsnLysAsnLysAsnSerLeuLeuIleGIuLeuSerThrGIuLeuPh 996
2401 AAAAATTAACAAAGAAATTCCTGTGATGGACCTTCAACTGGTCTGTT 2450
996 eAspAlaAsnAsnProLysMetLeuArgThrCysSerLeuProAspLeuS 1013
2451 TGATGCCAAACACCAACAGATGTTAAGACATGTTCCACTTCAGATCTCT 2500
1013 eTrpLysLeuPheArgThrLeuMetAspValIProThrValGIuAspValArg 1029
2501 CAAGGCTGTCAGAACCCCTTATGATGTTCCACCGTAGAGAGATGTTCT 2550
1030 GIuAspAsnLeuGIuIleAspGIuIleLysAspGIuAsnIleLysGIuGI 1046
2551 CAAGACAAATCTGAAATAGATGAATTAAGATGAACATTAAGAAAGAG 2600
1046 YProSerASpSerGIuAspIleValIlePheGIuLysThrAspThrAspLeu 1063
2601 ACCTTCGATCTGAAACACTTGTGTTGAAAGAACTGACACAGATTAC 2650
1063 InGIuLeuGIuAlaSerMetGIuGIuLeuLeuArgGIuGIuProGIuGIu 1079
2651 AAGAGCTGAGGCTCGATGAGACAGATCTATTAGGAAACACCTGTGTAA 2700
1080 GIuTrpSerGIuGIuGIuLysSerValLeuLysAsnSerAspValGIuTr 1096
2701 GAATACAGTGAAGAAGAGAGTCAGCTTGAAGAACAGATGTGGAGCC 2750
1096 oThrAlaAsnGIuThrAspValAlaAspGIuAspAspAsnProSerSerg 1113
2751 AACGCAAAATGGGACAGATCTGCGAGATGAAAGATGCCAATCCAGTATG 2800
1113 IuSerAlaLeuAsnGIuGIuTrpHisSerASpAsnSerASpGIuGIuIle 1129
2801 AAAGTCCCTGAACGAAGAAATGCGCACTCAGATTAACAGTATGTGAAT 2850
1130 AlAserGIuLysGIuLysAspSerValIlePheAsnHisLeuGIuLysLeu 1146
2851 GCTAGTAATGTGAATGCGATAGTGTCTTAAACATTTAGAGAGACTAG 2900
1146 GIeunHisLeuGIuGIuGIuMetGIuPheGIuLysPhePheGIuValIle 1163
2901 ACTTCATCTGAGACAGAAATGGGCTTTGAAAAATTCCTTGGAGTTTATG 2950
1163 IuLysIleLysAlaIleHisGIuAspGIuAspGIuAsnIleGIuLys 1179
2951 AGAAATTAAGGCTATTCAAGATGAAGATGAAGATTAATTAATTAATGT 3000
1180 SerLysIleValGIuAsnIleLeuGIuAsnGIuHisGIuHisLeuTrpAl 1196
3001 TCAAAAATAGTTCAAAATATTTTGGGAAATGAACATCACACATCTTATAG 3050
1196 aLysIleLeuHisLeuValMetAlaAspGIuAlaTrpGIuGIuLysAspAsn 1213
3051 CAAGATCTTCATTTAGTATGATGCGAGATGAGAGCTTAACAAAGATTAATG 3100
1213 spGIu 1214
3101 ATGAA 3105

```



```

317 eProLeuAlaTyrLysLysTyrGlyAspLysLysLeuHisGluLysLysP 334
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1526 GCCTTTAAATATTAAGAAATGAGATGAGATTAATAAAGCTTGAAGAAAAAC 1575
334 rOLeugInLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1576 CACCCCAATTAATTAACAGGCCATCAATTCCTGGAAGAAATGAAT 1625
351 ThrGluGluArgLysLysSerGluGluAlaAlaArgLysArg 367
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1626 TCTGAGAGAAAGAAAGAAATGCTGAGAGACGACGAAAAAAGAG 1675
367 gLeuGluPheIleGluLysGluLysGluLysAspGlnIleIleSerL 384
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1676 GTTGCAATTTTGAAGAAAGAAAGAAAGATCAG..ATTAGGT 1722
384 eumEtLysAlaGluGlnMetLysArgGlnLysGluArgLeuGluArg 400
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1723 TCTGGAAGGCTGAGCAGATGAGAGCGCAAGAGACGCGTTGGAGAG 1772
401 IleAsnArgAlaArgGluGlnGlyTyrPArgAsnValLeuSerAlaGly 417
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1773 ATAAATAGGGCCAGGAAACAAGATGAGAAATGTTTAAAGCGCTGCTGG 1822
417 ySerGlyValLysLysAlaProPheLeuGlySerGlyGlyThrIleAlaP 434
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1823 AAGCGGTGAAGTAAAGGCTTCCTTTTGGCATTTGAGAGCGGCTGCTCTC 1872
434 rOSerSerPheSerSerArgGlyValIntyGlnHisIleAlaIlePhe 450
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1873 CATCACCGCTTCTCTCGAGCGCCAGATGACATTACCATGCGCATTTT 1922
451 AspGlnMetGlnGlnGlnArgAlaGluAspAsnGluAlaLysTyrPArg 467
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1923 GACCAATGCCAGCGGCTAAGAGCAAGATTAATGAAGCAAGATGAAAGG 1972
467 gGluIleIleGlyArgLysLeuProGluArgLysGluLysGlnLeuAla 484
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1973 GGGAAATCATAGTGTGATGCTCCCAAGAAAGCAAAAGACCTTAGCTG 2022
484 aLIGluArgAlaLysGlnValGluGluPheLeuGlnArgLysArgGluAla 500
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2023 TAGAGAGAGCCACCAAGTGGAAGATTCCTACAGCTTAAACGAGAGCT 2072
501 MetGlnAsnLysAlaArgAlaGluGlyHisMetValTyrLeuAlaArgLe 517
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2073 ATGCAGATTAACCCCGAGCCGAAAGAGACACGTGTTATTTGGCAAGCT 2122
517 uArgGlnIleArgLeuGlnAsnPheAsnGluArgGlnGlnIleLysAlaL 534
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2123 GAGCAAAATAGACTACAAAATTTTATGAGCGCCAAACAGATTAAGCCA 2172
534 ySLeuarArgGlyLysLysGlnAlaAsnHisSerGluGlyGlnGluGly 550
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2173 AACTTCGGTGGAGATTAAGAGCTGATGTGCCAAAGCAAGCAAGAGCA 2222
551 SerGluGluAlaAspMetArgArgLysLysIleGluSerLeuLysAlaH 567
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2223 ACTGAAGAGACTGACATGAGGCTCAAAAAGATGAGTCACTTAGAGCGCA 2272
567 sAlaAsnAlaArgAlaValLeuLysGluGlnLeuGluArgLysArgL 584
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2273 AACCAATGACGCTGCTCTGCTACTAAAGAAAGAGCTGAGCGCAAGAA 2322
584 ySGluAlaTyrGluArgGluLysLysValTyrPGLuGlnHisLeuValAla 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2323 AGGAAGCTTATGAAGAAAGAAAGAAATGATGAAGAAACATTTGGCGGG 2372
601 LysGlyValLysSerSerAspValSerProProLeuGlyGlnHisLysL 617
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2373 AGG...GTAAAGAGCTCAGATGTTCTCTGCTTGAAGCTTCTTGAAG 2419
617 rGlyGlySerProSerLysGlnGlnMetArgSerValIleSerValThrS 634

```

```

2420 AGGTGGTTCCTCAACAAGCAGCAGGTGAAGCCTGCTCATTTCTGCACTT 2469
634 eAlaLeuLysGluValGlyValAspSerSerLeuThrAspThrArgLys 650
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2470 CAGCTTGAAGAAAGTGGCGCTGATGAGAAATTAAGTATGATCCACAGAA 2519
651 ThrSerGluGluMetGlnLysThrAsnAlaIleSerSerLysArgGly 667
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2520 .....GAAATAATGAAAAAGATTAACATGCTATTTCAGTAAGCGAGA 2563
667 uIleLeuArgArgLeuAsnGluAsnLeuLysAlaGlnGluAspGlyLys 684
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2564 AATCTGCTGAGCTTAATGAATGATTAAGCTCAAGAGATGAAAGG 2613
684 lMetGlnAsnLeuSerAspThrPheGluIleAsnValHisGluAspAla 700
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2614 AAAAGCAGCATCACTCAGTCTTGTGAGACCGTTGCTCACAAAGATGAG 2653
701 LysGluHisGluLysGluLysSerValSerSerAspArgLysLysTyrPGL 717
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2664 AGAGAGTATGAGACAGAAATGCCATTTCTCTGATCCCAAGAACTGGGA 2713
717 uAlaGlyGlyGlnLeuValIleProLeuAspGluLeuThrLeuAspThrS 734
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2714 GATGGAGGCTCAGCTTGTGATTCCTCTCGATGACAGTACAGTATACAT 2763
734 eRPheSerThrThrGluArgHisThrValGlyValIleLysLeuGly 750
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2764 CCTTCTCTGCAACGAAAAACATACCTGTGGAGAGATTATTAATTAAGAT 2813
751 ProAsnGlySerProArgArgAlaTyrPGLysSerProThrAspSerVal 767
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2814 TCTAATGCTCTCCAGAAAGATCTGGGGAAGAAACCCACAGATTCCT 2863
767 lLeuLysIleLeuGlyGluAlaGluLeuGlnLeuGlnThrArgLeuLeuG 784
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2864 GCTGAATATCTTGAAAGCTGAATTAACAGCTATAGACAGACACTAGG 2913
784 lAsnThrThrIleArgSerGluIleSerProGluGlyLysLysTyrLys 800
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2914 AAAACACATCTTTTAAAGTGAAGTTATGCTGAAGAGAGAGACATCAAA 2963
801 ProLeuIleThrGlyLysLysValGlnCysIleSerHisGluIleAs 817
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2964 CCTTACTACTGAGAGAGAGATTCGACGTCCATTTCAAAAGAAATATAA 3013
817 nProSerAlaIleValAspSerProValGluThrLysSerProGluPheS 834
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3014 TCCATCAGCTACCTGTGATTC...ACTGAAGCAAGAAAGTCCAAAGTTTA 3060
834 eArgAlaAspSerProGlnMetSerLeuLysLeuGlyLysAsnLeuGluL 850
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3061 CTGAGGCTCTCCCAAAATGTCA.....GAAAGAAATGTGGAAGAA 3101
851 ProAspAspLeuGlnThrGluIleLeuGlnGluProSerGlyThrAsnL 867
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3102 CCTGATGATTTGGAACAGAGTCTACAAAGAGCCAAAGTACAGACACAC 3151
867 sAspGluSerLeuProCysThrIleThrAspValTyrPLeuSerGluL 884
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3152 AGATGGAGATTGCCACTGTTCTTATATGATGTGTGACTACAGAGAAAG 3201
884 ySGluThrLysGluThrGlnSerAlaAspArgIleThrIleGlnGluAsn 900
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3202 AAGAGCTTAAGAAAGTGAAGTGAAGAAAGTGTGCTGTGACAGAGAT 3251
901 GluValSerGluAspGlyValSerSerThrValAspGlnLeuSerAspI 917
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3252 GAAGTTGTGAAGATGAATTCAGAGCAAGCTGACCAATCCTGTAAAGA 3301
917 eHisIleGluProGlyThrAsnAspSerGlnHisSerLysCysAspValA 934
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

3302 TCAGAGAGATCTCGAGTAGAGACGATTCGCCAGTCTGGCTGTGATGTAG 3351
934 splysserValgInprogluProphenhIslyValaHisSerGlu 950
3352 AGAAGTCATACAGCCAGATTCCTCCAGAAAGTGTTCAATTCATTCAG 3401
951 HisLeuAspLeuValProgluValgInserValgIncysSerProglu 967
3402 GACTGTGACTTA.....GTTGAGCAGATTCCTCCACGAGAGA 3442
967 userPheAlaPheArgSerHisSerHisLeuProProlysAnlysAnL 984
3443 ACCAATTCCAATTCGATCTCAGTCTGATTCCTCCACCAAACTAAGACA 3492
984 AsnSerLeuLeuIleGlyLeuSerThyGlyLeuPheAspAlaAsn 1000
3493 AGAATTCCTTACTGATTCGATTCCTCCAGGTCCTGTTGATGACAACAT 3542
1001 ProlyMetLeuArgThyCysSerLeuProAspLeuSerLyLeuPheAr 1017
3543 CCAGAGATGCTGAGACCTGCTCCTCAGATCTTCCTCCAGCTGTTGAG 3592
1017 gThrLeuMetAspValProThyValgIlyAspValArgInAspAnLeu 1034
3593 AACCTTATGACGCTCCACGCTGCGAGCGATTCACAGACAGCTCTTG 3642
1034 IuileaspGluIleLysAspGluAsnIleLysGluGlyProSerAspSer 1050
3643 AAATCGATGAGCTGGAAGATGAACCAATTAAAGAGGCTCTCTATTCC 3692
1051 GluAspIleValPheGluGluThrAspThrAspLeuGluInLeuGlnAl 1067
3693 GAGGACACTGATTATTGAAAGAACTGACACAGATTTCACAGACCTTCAG 3742
1067 aSerMetGluInLeuLeuArgGluInProgluGluGluThySerGlu 1084
3743 CTCATGAGACGCTGCTAGGAGCAACAGCTGACGATACATACAGTAGAG 3792
1084 IuGluGluSerValLeuLysAsnSerAspValgIuProThrAlaAsnGly 1100
3793 AGGAGAGCTGTTTAAAGCGAGCGATGTGAGCAGACAGACAGAGAGGG 3842
1101 ThrAspAlaIAspGluAspAspAsnProSerSerGluSerAlaLeuAs 1117
3843 ACAGATGCCCGACGAGGAGGACACCCAGCAGGAGAAAGGCC..CTGAA 3891
1117 nGluGluThrPheHisSerAspAsnSerAspGlyIuIleAlaSerGluCysG 1134
3892 CGAGAGATGGCACTACATATATAGTGCCTGAGACACACAGATGATGG 3941
1134 IucysAspSerValPheAsnHisLeuGluGluLeuArgLeuHisLeuGlu 1150
3942 AATATGACGTGCTTAAACATTTAGAGAACTAGACATTCCTGCTGGAG 3991
1151 GIngluMetGlyPheGluLysPhePheGluValThyGluLysIleLysAl 1167
3992 CAGAAATGGCTTTGAAAGTTCTTTGAGGTTTATGAAATTAAGAGG 4041
1167 alLeuHisGluAspGluAspGluAsnIleGluLecysSerLysIleValg 1184
4042 TATTCATGAGATGAAGATGAATAATTGAAATTTGTTCAACAATGTTG 4091
1184 InAsnIleLeuGlyAsnGluHisGlnHisLeuThyAlaLysIleLeuHis 1200
4092 AGAATATTTTGGCAATGAGCCAGCATCTCTATGCCAACAAATTCGCAT 4141
1201 LeuValMetAlaAspGlyAlaThyGlnGluLysAspAsnAspGlu 1214
4142 TTATGATGCGAGATGAGACCTATCAGAAAGATTAATGATGAA 4183

seq_name: gp_pr:AF155113
seq_documentation_block:

```

```

LOCUS AF155113 3555 bp mRNA linear PRI 05-JAN-2000
DEFINITION Homo sapiens NY-REN-55 antigen mRNA, partial cds.
ACCESSION AF155113
VERSION AF155113.1 GI:5360120
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 3555)
Scanlan,M.J., Gordon,J.D., Williamson,B., Stockert,E., Bander,N.H.,
Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T.
and Old,L.J.
Antigens recognized by autologous antibody in patients with
renal-cell carcinoma
Int. J. Cancer 83 (4), 456-464 (1999)
JOURNAL
Int. J. Cancer 83 (4), 456-464 (1999)
MEDLINE
10508479
PUBMED
2 (bases 1 to 3555)
Scanlan,M.J., Gordon,J.D., Williamson,B., Stockert,E., Bander,N.H.,
Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T.
and Old,L.J.
Direct Submission
Submitted (28-MAY-1999) Ludwig Institute, Sloan-Kettering
Institute, 1275 York Ave, New York, NY 10021, USA
FEATURES
source
1..3555
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="1973/10-4"
/cell_type="renal cell carcinoma"
<..2317
/note="similar to Mus musculus NRK serine/threonine
kinase"
/codon_start=2
/product="NY-REN-55 antigen"
/protein_id="AAD42879.1"
/db_xref="GI:5360121"
/translation="MHYATFDOMQOARADNEAKRKREIYGRGLPERQKGLAVERA
KQVERPQORPRAMONKARAGHWYLLARQIOLQFNPROQIKAKLGEKKENHS
EKGESFEDAMERKRETSIKAHANRAAVLKEQERKRAYREKKWEHLEIVAKGV
KSDVSPPLGCHRENGSPSKOMNSVTSYSALEKEVDLSLDTRETSSEMQKTNA
ISKRILRLRLNEINAKQEDKRONISDPEINVEDDAKHEKRSVSDRRKNEAG
GQVLPLDELTLIDSPFTEBHRTGQYIKIGPNSSPRRAGKSPDVSUKIGAEILO
LOVELLENTINSEISPRGEKYLIDEPGCTKRKSDLSCTTDVWISBEKRETSAD
ASPOMSLKLEGNLEEDDELEIILDEPGLNDSQHSKCPDYSKQVQPEPRHVVSEH
RTIQENESVEDSVSDYDQSDIHLEPGLNDSQHSKCPDYSKQVQPEPRHVVSEH
LNLVPOVSQCSPESEAFRSHSLPEKKNKNSLIGSLGLPANNPKMLPTCSL
PDLKLRITIMVPTVGVRODNLIDELDENKRESDSDSDIYPERPDITLOFQIA
SMOILREDPGEYSSEESVLKNSDVETPANGTGDVDEDDNFSSEALNENHSDNS
DGEIASECCDSEVFNHLELRHLEQENGFEFEYEKIKALIHEDEDENIEISKIY
QNLLENHOLHYAKLILHVMADGAYQEDNDE"
BASE COUNT 1247 a 572 c 728 g 1008 t
ORIGIN
alignment_scores:
Quality: 3945.00 Length: 771
Ratio: 5.130 Gaps: 0
Percent Similarity: 99.741 Percent Identity: 99.611
alignment_block:
US-09-783-320-4 x AF155113
Align seg 1/1 to: AF155113 from: 1 to: 3555
444 GlnHisThyHisAlaIlePheAspGluMetGluInGlnGluArgAlaGluAs 460
|||||
2 GAACATTTACAGGCCATTTTTCACCAATTCGACGACCAAGAGAGAGA 51
|||||
460 PaangluAlaLysTrpLysArgGluIleThyGlyArgGlyLeuProgluA 477
|||||
52 TAATGAAGCTAAATGGAAGAAATATATGCTGAGGCTCTCCAGAAA 101
|||||

```

477 rglInlysglInleuAlaValgluArGAlaLysglInValgluIuphe 493
 102 GGCAAAAGGCGACGTAGCTGTAAAGAGCTTAACAGTAGAAGAGTTC 151
 494 leuGlnArGlyArGluAlaMetGlnAsnLysAlaArGAlaIuIyH 510
 152 CTGCAGCGCAAAACGGAAGCTATGCAGATTAAGCTGACCCGACAGACA 201
 510 sMeValIyrtleuAlaArGleuArGInIleArGleuGlnAsnPhaAsnG 527
 202 TATGTTTATCTGGCAAGACAGAGCAAAATTAACACTACAGAAATTTCAAG 251
 527 luArGglInGlnIleLysAlaLysleuArGlygluLysgluAlaAsn 543
 252 AGCGCCAGACAGATTAAAGCCCAAACTGTGTGTAAGAAAGAAAGCTAAT 301
 544 HisSerGluGluGluGluGluSerGluGluAlaAspMetArGArGly 560
 302 CATCTGAAAGAGCAAGAAAGAGAGAGCTGACATGAGCGCAAAAA 351
 560 sIlleGluSerleuLysAlaHisAlaAsnAlaArGAlaAlaValleuLysG 577
 352 AATCGAAATCACTGAAGAGCCCATGCMAATGCAGCTGCTGTACTTAAAG 401
 577 luGlnleuGluArGlyArGlysgluAlaTyrgluArGlyLysVal 593
 402 AACAACTGAAAGCAAGAAAGAGAGGCTTATGAGAGAGAAAAAAGTG 451
 594 TrpGluGluHisleuValAlaLysGlyValLysSerSerAspValSerP 610
 452 TGGGAAGAGCATTTGTGGCTTAAAGAGTTAAGATTCTGATTTCTCC 501
 610 oProleuGlyGlnHisGluTrpGlyLysSerProSerLysGlnIleMetA 627
 502 ACCTTGGGACAGCATGAACAGGTGGCTCTCCATCAACCAACAACTGA 551
 627 rGSerValIleSerValThrSerAlaLeuLysGluValGlyValAspSer 643
 552 GATCTGATTTCTGTACTCTACCTTGAAGAGAGTTGGCTGGACAG 601
 644 SerleuThrAspThrArGluThrSerGluGluMetGlnLysThrAsnS 660
 602 AGTTTAACCTGATACCCGCGAAACTTCAGAAAGATGCMAAAGCAACAA 651
 660 nAlaIleSerSerLysArGluIleleuArGArGleuAsnGluAsnleu 677
 652 TGCTATTTCAGTAAGCGAGAAATACCTCGTAGATTAAATGAATAACTTA 701
 677 ySAlaGluGluAspGluLysGlyMetGlnAsnleuSerAspThrPheGlu 693
 702 AAGCTCAAGAGATGAAAGAAAGAAAGCAAACTCTCTGATCTTTGAG 751
 694 IleAsnValHisGluAspAlaLysGluHisGluLysGluSerValSe 710
 752 ATAAATGTTTCATGAAGATGCCAAAGACATGAAAAAAGAAAAACACTTTC 801
 710 rSerAspArGlyLysTrpGluAlaGlygluInleuValIleProleuA 727
 802 ATCTGATGCAAGAGTGGAGCGAGAGCTTGTGATTCCTCTGCG 851
 727 sPgluLeuThrleuAspThrSerPheSerThrGluArGHisThrVal 743
 852 ATGAGTTAACACTAGATACATCTCTCTCACTGAAAGACATTAACAGAG 901
 744 GlyGluValIleLysleuGlyProAsnGlySerProArGArGAlaTrpGl 760
 902 GGAGAGAGTTATTAAATAGTCCATATGATCTCCAAAGAGAGACCTGGGG 951
 760 yLysSerProThrAspSerValleuLysIleleuGlyGluAlaGluLeuG 777
 952 GAAAGTCCGACAGATTCTGTCTTAAAGATACTTGGAGAGCTGCAACTAC 1001

777 InleuGlnThrGluLeuLeuGluAsnThrIleArGSerGluIleSer 793
 1002 AACTTCAGACAGCAATATTAGAAAAATACAACTATTAGAAAGTAGATTCT 1051
 794 ProGluGlygluLysTrpLysProleuIleThrGlygluLysValGI 810
 1052 CCGGAAGGGGAAAAAGTACAACCCCTTAATTACTGGAGAAAAAAGTACA 1101
 810 nCysIleSerHisGluIleAsnProSerAlaIleValAspSerProVAG 827
 1102 ATGTATTTCATGAATATAACCATACACTTGTGTGATTCCTCTGTTG 1151
 827 LuThrLysSerProGluPheSerGluAlaSerProGlnMetSerleuLys 843
 1152 AGACAAAAGTCCCGAGTTCAAGAGCAATCTCCACAGATGCATTGANA 1201
 844 leuGluGlyAsnleuGluGluProAspAspLeuGluThrGluIleleuGI 860
 1202 CTGGAAGGAAATTTAGAAAGACCTGATGATTTGAAAGCAAGAAATTTACA 1251
 860 nGluProSerGlyThrAsnLysAspGluSerleuProCysThrIleThra 877
 1252 AGAGCCAAAGTGAACAAAGAGATGAGCTTGCATGCACTATTTACTG 1301
 877 sPValTrpIleSerGluGluLysGluThrLysGluThrGlnSerAlaAsp 893
 1302 ATGTGTGATTACTGAGGAAAAAGAAACAAAGAAACACAGTCGCGCAGAT 1351
 894 ArgIleThrIleGlnGluAsnGluValSerGluAspGlyValSerSerTh 910
 1352 AGGATCAACATTCAGGAAATGAAATGATTTCTGAAGATGAGTCTCGAGTAC 1401
 910 rValAspGlnleuSerAspIleHisIleGluProGlyThrAsnAspSerG 927
 1402 TGTGGACCAACTTATGACATTCATATAGAGCTGGAAACCAATGATTCTC 1451
 927 InHisSerLysCysAspValAspLysSerValGlnProGluProPhePhe 943
 1452 AGCACTCTTAATGTGATGTAGATAGCTGTGTCAACCGGAACCATTTTTC 1501
 944 HisLysValValHisSerGluHisleuAsnleuValProGlnValGlnSe 960
 1502 CATAAAGTGGTTTCATTTCTGACACCTTGAACTTAGTCCCTCAAGTTCAATC 1551
 960 rValGlnCysSerProGluGluSerPheAlaPheArGSerHisSerHisL 977
 1552 AGTTCACTGTTCACCAAGAAATCCCTTTCATTTGATCTCACTCGCATTT 1601
 977 euProProLysAsnLysAsnLysAsnSerleuLeuIleGlyLeuSerThr 993
 1602 TACCACCAAAAAATTAATAAACAGAAATTCCTGCTGATTTGACTTCAACT 1651
 994 GlYleuPheAspAlaAsnAsnProLysMetleuArGThrCysSerleuPr 1010
 1652 GGTCTGTTTCATGCAAAACCAACCAAGATGTTAAAGACATGTCTCTCC 1701
 1010 oAspLeuSerLysleuPheArGThrleuMetAspValProThrValGlyA 1027
 1702 AGATCTCTCAAGCTGTGTGAGACCTTATGATGTCTCCACAGTAGAG 1751
 1027 sPValArGlnAspAsnleuGluIleAspGluIleLysAspGluAsnIle 1043
 1752 ATGTTCTCAAGACAATCTTGAATATGATGAATTAAGATGAAGAAAAATTT 1801
 1044 LysGluGlyProSerAspSerGluAspIleValPheGluGluThrAspTh 1060
 1802 AAAAGAGAGACCTTCTGATTTCTGAAAGACATTTGTTTGAAGAAACGACAC 1851
 1060 rAspLeuGlnGluLeuGlnAlaSerMetGluInleuLeuArGlyGluGlnP 1077
 1852 AGATTTCAGAGAGCTGAGGCTTCGATGGAACAGTTACTTAGGGAACAC 1901
 1077 roGlyGluGluLysSerGluGluGluSerValleuLysAsnSerAsp 1093

```

|||||
1902 CTGCTGAGATACAGTGAAGAGAGTCTGTTGAGACAGTAT 1951
1094 ValGIPrThrAlaAsnGlyThrasPalaAlaAspGluAspAsnPr 1110
1952 GTGAGCCAACTGCAATGGACAGATGTCGACATGAATGACATCC 2001
1110 oSerSerGluSerAlaLeuAsnGluGluTrpHisSerAspAsnSerasp 1127
2002 CAGCAGGAGAAAGTCCCTGACAGAGAAATGGACATCAATACAGTGTG 2051
1127 LylGluLeuAlaSerGluGluGluGluGluGluGluGluGluGlu 1143
2052 GTGAATGCTGAGTGAATGAGATGAGTGTCTTTTACCATTAGAG 2101
1144 GluLeuArgLeuHisLeuGluGluGluGluGluGluGluGluGlu 1160
2102 GAACCTGAGCTCATCTCGAGCAGGAATGGCTTTGAAATTCCTTGA 2151
1160 uValTyrGluLysIleLysAlaIleHisGluAspGluAspGluAsnIleG 1177
2152 GGTATTAGAGAAATAAGCTATTCATGAGATGAAGTGAATATTTG 2201
1177 LuIleCysSerLysIleValGlnAsnIleLeuGluYAsnGluHis 1193
2202 AAATTTGTCAAAATAGTCCAAAATATTTTGGAAATGCAACATCATG 2251
1194 LeuTyrAlaLysIleLeuHisLeuValMetAlaAspGluYalaTyrGln 1210
2252 CTTTATGCGCAAGATTCCTTCATTAGTCATGCGAGATGAGCCTTACC 2301
1210 uAspAsnAspGlu 1214
2302 AGATATATGATGAA 2314

seq_name: gb_AK025658
seq_documentation block:
LOCUS AK025658 2466 bp mRNA linear PRI 29-SEP-2000
DEFINITION Homo sapiens cDNA: FLJ22005 fis, clone HEP06902.
ACCESSION AK025658
VERSION AK025658.1 GI:10438246
KEYWORDS oligo capping, fis (full insert sequence).
SOURCE Homo sapiens hepatoma cell_line:hep02 cDNA to mRNA, clone_11b:HEP
clone:HEP06902.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Ohtani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEBO human cDNA sequencing project
Unpublished (2000)
2 (bases 1 to 2466)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT
NEBO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology: cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
1..2466
location/Qualifiers
source
/organism="Homo sapiens"

```

```

/ab_xref="taxon:9606"
/clone="HEP06902"
/cell_line="hep2"
/cell_type="hepatoma"
/clone_id="hep"
/note="cloning vector pME18SFL3"
60..1187
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15207.1"
/ab_xref="GI:10438247"
translation="MSLKLEGNLEPEDDLEFTELLOPSTKNDSELPCTITDVFISE
KETKETSDADRTTQENVEDSVSTVQSLDTHPPGTDSDHSDSKDNDKSVQPP
FEHRYVHSEHLNLPVOYVOCSPESFAFRSHSLPRKKNKNSLIGLSTGFLDN
NPKMLRTGSLPDLSEKLRITLMDYPTVGDRKDLNLEDEENIKKEPSSEDIYFE
TDDTDLQASMEQLREQPEYSEESVYLNKSDVEPTANGTVADEDDNPSSEA
LNEWHSDNSDGEIASECEDSVFNLLEIRLRLDEQGEKFEFEYERITAIHEDD
ENIECSKIVONILIGNHOLHYAKRLHLVADGAYQENDF"
BASE COUNT 857 a 382 c 465 g 762 t
ORIGIN
alignment_scores:
Quality: 2048.00 Length: 394
Ratio: 5.198 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.746
alignment_block:
US-09-783-320-4 x AK025658 ..
Align seg 1/1 to: AK025658 from: 1 to: 2466
821 llevalasppserprovalgluthrlysserproglupheserglualse 837
|||||
3 attggtgattctccgttgagacaaaagtcggaatttcagagagcattc 52
837 rproglmetserleuylsleugluclyasnlengluinuproaspapl 854
|||||
53 tccacagatgcatgtaagaaactggaagaaatttagaagaaactgatt 102
854 eugluthrglulleuglucgluproserglythrnslyssaspgluser 870
|||||
103 tggaaacagaaaattctacaaagcgaagtgaaacaaacaaatgagacc 152
871 leuprocysthrilethrpsvaltrpileserglucglulysgluthry 887
|||||
153 ttgccatgcactattactgatgtgattagtagagaaacaaacaaacaa 202
887 sgluthrglnseralaasparglylethrileglnuasnglualse 904
|||||
203 gcaaaactcagtcgacagatagatcacattcagaaacaaacaaacaa 252
904 luspalyalsersethrvalaspglnleuseraplienisilegu 920
|||||
253 aacatgagctcgtcagatcgtgacacaaacttactgacattcatatag 302
921 progltythrpsasercglnhisserlyscysasppalaaslysserya 937
|||||
303 cctggaaccaaagattctcagacacttaaatgattatagattagctgt 352
937 lginuprogiluprophehislyvalvalhisserglnhisleuasnl 954
|||||
353 gcaacggaaacatttttcapaaagtggttcatttcgaacacttgaa 402
954 euvalproglunvalinservaylincysserproglugluserphala 970
|||||
403 tagtccctcagaattcaatcagttcagtttcacacgaagaacattgca 452
971 pheargserthserthleuuproprolysasnllysaasnlserle 987
|||||
453 ttctgattctcagtcgattttaccacaaacaaacaaacaaacaaatg 502
987 uleuileglyleuserthrlyleupheaspalaasnsnprolysmetl 1004

```

```

|||||
503 GCTGATGAGACTTTCACACTGCTGTTGATGCAACACCAACCCAAATGATG 552
1004 eUaATThCysSerLeuProAspLeuSerIysLeuPheArgThrLeuMet 1020
553 TAAAGACATGTTCACTTCACATCTCTCAAGCGTTCACAAACCCCTTAA 602
1021 AspValProThrValGlyAspValArgGlnAspAsnLeuGluIleAspGly 1037
603 GATGTTCCACCGTAGAGATGTCGTCAAGACAACTTGAATTAATATGATGA 652
1037 uIleIysAspGluAsnIleIysGluGlyProSerAspSerGluAspIleVal 1054
653 AATTCAAGATGAAAACATTAAAGAAAGACCTTCTGATTCGAAAGACATTC 702
1054 aPheGluGluThrAspThrAspLeuGluLeuGlnIleAlaSerMetGlu 1070
703 TGTTTGAAGAAACTGACACAGATTTCACAGAGCTGCAGCGCTCGATGGA 752
1071 GlnLeuLeuArgGluGlnProGlyGluGluIleIysSerGluGluIle 1087
753 CAGTTACTTAGGGAACAACCTGCTGGAAGAAATACAGTGAAGACAGACAGTC 802
1087 rValLeuIysAsnSerAspValGluProThrAlaAsnGlyThrAspValAla 1104
803 AGCTTTGAAGAAACAGTGTGTGGAGCCCACTGCAAAATGGGACAGATGCG 852
1104 laAspGluAspAspAsnProSerSerGluSerAlaLeuAsnGluGluTrp 1120
853 CAGATGAGATGACAAATCCAGCAGTGAAGTGCCTCGAAGCAAGAAATCG 902
1121 HisSerAspAsnSerAspGlyGluIleAlaSerGluGluGluGluAspSer 1137
903 CACTCAGATTAACAGTGAATGCTGAAGAAATGCTAGTGAATGCAATGCCATAG 952
1137 rValPheAsnHisLeuGluGluLeuArgLeuHisLeuGluGluMetG 1154
953 TGTCTTTAACCTTTAGAGAACTGAGACCTTCATCTGAGACAGAAATGG 1002
1154 lYpHeGluLysPhePheGluValIleGluIleIysAlaIleHisGlu 1170
1003 GCTTTGAAAATCTCTTGAGGTTATGCAAAATAAGGCTATTCATGAA 1052
1171 AspGluAspGluAsnIleGluIleCysSerIysIleValGlnAsnIleLe 1187
1053 GATGAAGATGAATAATGAAATTTGTTCAAAATACTTCAAAATATATTTT 1102
1187 uGlyAsnGluHisGlnHisLeuTyralAlaIysIleLeuHisLeuValMet 1204
1103 GGGAAATGACATCAGCATCTTTATGCCAAGATCTTCATTTAGTCATCG 1152
1204 laAspGlyAlaIleTyrglnGluAspAsnAspGlu 1214
1153 CAGATGGAGCCTACCAAGACATATGATGAA 1184

```

seq_name: gb_pr:AK027580

seq_documentation_block:

LOCUS AK027580 2975 bp mRNA linear PRI 15-MAY-2001

DEFINITION Homo sapiens CDNA FLJ14674 fls, clone NT2RP2003912, highly similar

TO SERINE/THREONINE-PROTEIN KINASE NEK1 (EC 2.7.1.-).

ACCESSION AK027580.1 GI:14042352

VERSION AK027580.1

KEYWORDS oligo capping; fls (full insert sequence);

SOURCE Homo sapiens fetalocarcinoma cell, line:NT2 CDNA to mRNA,

clone_lib:NT2RP2 clone:NT2RP2003912.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (sites) Isegai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,

```

Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuno,Y. and Sasaki,N.
MEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2975)
Isegai,T. and Otsuki,T.
Direct Submission
Submitted (10-MAY-2001) Takao Isegai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yama, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hrl.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
MEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; CDNA full insert sequencing;
Research Association for Biotechnology; CDNA library construction,
5'-8'3'-end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source
Location/Qualifiers
1..2975
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP2003912"
/cell_line="NT2"
/cell_type="fetalocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
354..1316
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB5209.1"
/db_xref="GI:14042353"
/translation="MKRLVLRITSGSEFPVSLHYSDLRSLVSLQFKRNPDRPSVNS
ILEGFTAKRIERKFLSPQLIAEERCFKTSKGSOPIPAKRPASGONSISVWPAOKIT
KPAKVGIPLAIVKYGDKKLHEKKPIQKHOKHONPEKRVNTEGERRKISEGAARRK
LEPTEKKOKDOITSLMKAEOMKROEKRLERINARPCGVRNVLSAGSGSEYKAPR
LGSGCTAPSPSFSRGCEYHYHAIPDOMOQOARDEMAKREIYIGRLPERGLIPGV
RPGPYRAGHHHPDMDIRTKLRKLKAVSKOANNRIDCDP"
BASE COUNT 894 a 556 c 562 g 963 t
ORIGIN
alignment_scores:
Quality: 2012.00 Length: 387
Ratio: 5.199 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-783-320-4 x AK027580 ..
Align seg 1/1 to: AK027580 from: 1 to: 2975
91 ArgTlleAsnAlaGlnIlysgIyValleupheGngInluAspGlnIleLeuAs 107
|||||
3 CGAATAAATGCTCAGAAAGCGTTTGTTCAGAGATCGATCGATTTTGA 52
107 pTppheValGlnIleCysLeuAlaLeuIyshIshAspArgIyxi 124
|||||
53 CTGGTTTGACAGATATGTTGGCCCTGAAACATGATCATGATGAAAA 102
124 leIeuHisArgAspIleIySerGlnAsnIlePheLeuThrIlyAspGly 140
|||||
103 TTCTTCATCGACACATTAATCTCAGAAACATATTTTAACTAAAGATGGA 152
141 ThrValGlnLeuGluIlyAspPheGlyIleAlaArgValleuAsnSerThrVa 157
|||||
153 ACAGTACAACTTGAGATTTTGAATTTGCTAAGATTTCTTAATAGACTGT 202
157 lGluLeuAlaArgThrCysIleGlyThrProTyTrIyLeuSerProGluI 174
|||||
203 AGAGCTGGCTCGAAACTTGATAGAGGACCCCATACTACTTGACACTGAAA 252
174 lCysGluAsnIlyProTyTrAsnAsnIlySerAspIleTrpAlaLeuGly 190

```

```

|||||
253 TTCGTAAGAAACCTTACATATATAAAGTGACATTGGCTGGGG 302
191 CysValLeuTyrGluLeuCysThrLeuIleuHisAlaPheGluAlaGly 207
303 TGTCCTTATGACCTGTGTAACCTTAAACATCTTTGACGTGGAG 352
207 rMetLysAsnLeuValLeuLysIleIleSerGlySerPheProVal 224
353 TATGAAATCCCTGACTGATGATATATCTGATCTTTCCACCTGTCT 402
224 erLeuHisTyrSerTyrAspLeuAgsSerLeuValSerGluLeuPheLys 240
403 CTTTGATATATCTCTATGATCTCCCACTGGGTGGTCTGACATTATTA 452
241 ArgAsnProArgAspArgProSerValAsnSerIleLeuGluLysGly 257
453 AGAAATCCCTAGGATGACCATCATCACTCCATATTGGAGAAAGGTT 502
257 eIleAlaLysArgIleGluLysPheLeuSerProGluLeuAlaGlu 274
503 TATAGCCAAACCGATTGAAAGTTCTCTCTCTCAGCTTATTCAGAG 552
274 LuPheCysLeuLysThrPheSerLysPheGlySerGluProIleProAla 290
553 AATTGTCTTAAACATTTTCAGATTGATGATCAACCCATATCCAGCT 602
291 LysArgProAlaSerGlyGlnAsnSerIleSerValMetProAlaGlu 307
603 AAAAGCCAGCTTCAGAGCAAACTCGATTCTGTTATCCCTGCTCAGAA 652
307 sIleThrLysProAlaAlaLysTyrGlyIleProLeuAlaTyrLysLys 324
653 AATTACAAAGCTCCGCTTAATATGAAATACCTTACATTAAGAAAT 702
324 YRGLYAspLysLysLeuHisGluLysLysProLeuGluLysHisLysGlu 340
703 ATGGAGATATAAAATTTACAGAAAGAAAGCACTGCCAAACATATAAG 752
341 AlaHisGluThrProGluLysArgValAsnThrGluGluArgArgGly 357
753 GCCCATCAACTCCAGAGAAAGATGATATCTGAGAAAGAAAGGAGAA 802
357 sIleSerGluGluAlaAlaLysArgArgLeuGluPheIleGluLysG 374
803 AATATCTGAGAGCGAGAGAAAGAGAGAGCTGGAAATTTATTGAAAGG 852
374 LuLysLysGluLysAspGlnIleLeuSerLeuMetLysAlaGluMet 390
853 AAAAGCAACAAAGATCAGATTATATGATTATGAAAGGCTGACAAATG 902
391 LysArgGlnGluLysGluArgLeuGluArgIleAsnArgAlaArgGlu 407
903 AAAAGCAACAAAGAAAGAGCTTGGAAATATAATAGGCGCAGGAGACA 952
407 nGlyTyrArgAsnValLeuSerAlaGlyLysGlyGluValLysAla 424
953 AGGATGAGAAATGTGCTAGCTGCTGGAGAGTGAAGTGAAGTAAAGCTC 1002
424 rPheLeuGlySerGlyLysThrIleAlaProSerSerPheSerSerArg 440
1003 CTTTCTGGGCACTGAGGAGATATAGCTCATCTTTCTTCTCGA 1052
441 GLYGluTyrGluHisTyrHisAlaIlePheAspGlnMetGlnGlnIle 457
1053 GGCAGATATGACATTAACATCCATTTTTCACCAATGACAGCAACAAAG 1102
457 gAlaGluAspAsnGluAlaLysTyrLysArgGluIleTyrGlyArgLys 474
1103 AGCGAAGATATATGAAAGTAAATGAGAAAGAAATATATATGTCGAGCTC 1152
474 euProGluArg 477
|||||

```

```

1153 TTCGTAAGAAAGA 1163
seq_name: gb_pat:AX166554
seq_documentation_block:
LOCUS AX166554 1947 bp DNA Linear PAT 22-JUN-2001
DEFINITION Sequence 45 from Patent WO0138503.
ACCESSION AX166554
VERSION AX166554.1 GI:14546899
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1947)
Plowman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R.,
Flanagan, P. and Clardy, D.S.
Novel human protein kinases and protein kinase-like enzymes
Patent: WO 0138503-A 45 31-MAY-2001;
JOURNAL
Sugen, Inc. (US)
FEATURES
location/Qualifiers
source 1..1947
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 638 a 357 c 455 g 497 t
ORIGIN
alignment_scores:
Quality: 1025.00 Length: 818
Ratio: 2.293 Gaps: 25
Percent Similarity: 54.645 Percent Identity: 32.763
alignment_block:
US-09-783-320-4 x AX166554
Align seg 1/1 to: AX166554 from: 1 to: 1947
1 MetGluLysTyrValArgLeuGlnLysIleGlyGluGlySerPheGly 17
|||||
1 ATGGATTAAGTACGATGATTAAGCCATCGGCAAGCTTCGGGAA 50
|||||
17 sAlaIleLeuValLysSerThrGluAspGlyArgGlnTyrValIleLysG 34
|||||
51 ACCATACCTTAGCTAAAGGAGATCAGATACAGAAAGAGAGCTTCAAGAGA 100
|||||
34 LuIleAsnIleSerArgMetSerSerLysGluArgGluLysArgArg 50
|||||
101 AGATCAATTTTGAAGAAAGATGCCCATACAGAAAGAGAGCTTCAAGAGA 150
|||||
51 GluValAlaValLeuAlaAsnMetLysHisProAsnIleValGluTyr 67
|||||
151 GAAGTGATTTCTCTGAAAGATGAAACATCCACACTTGTACCTTCT 200
|||||
67 gGluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCysG 84
|||||
201 CATTCATTTCAAGAGAAATGCGAGCGCTGTTATGTGAATGAAATATGTC 250
|||||
84 LuGluLysAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
|||||
251 ATGAGAGGAGATCTCATGAAAGAGATCATAGACAACGCGGTGTTATTT 300
|||||
101 GlnGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeu 117
|||||
301 AGTGAAATTCAGATCTCGCTGGTGTGTACAGATTTCTAGAGACTAAA 350
|||||
117 sHisValHisAspArgLysIleLeuHisArgAspIleLysSerGluAsn 134
|||||
351 ACATATTCACAGAGAGATATTTACACAGGACATAAAGCTCAGAGACA 400
|||||
134 LePheLeuThrLysAspGlyThrVal...GlnLeuGlyAspPheGlyLys 149
|||||
401 TTTTCTTAGCAAGAGAGATGTGGCAAGCTTGGGACTTGGTATA 450
|||||

```

```

150 AlaAlaValAlaAsnSerThrValGluLeuAlaArgHisCysIleGlyThr 166
451 GGAGAGAGTCTCTAAATAAATTTCCATGGAACTCTCTGCAACTGTATATGGAC 500
166 rProGlyrThrLeuSerProGluIleCysGluAsnLysProGlyrAsnAsnL 183
501 ACCCTAACAACCTGTGCCAGAAATCTGTCAAAATAAACCCTACAAACATA 550
183 ySerAspIleThrPheAlaLeuGluCysValIleGlyrGluLeuCysThrLeu 199
551 AAACGAGATTTTGGCTCTGTGGCTGTGCTGTATATAGACTGTGCACACT 600
200 LysHisAlaPheGluAlaGluSerMetLysAsnLeuValLeuIleIleI 216
601 AAACATCCTTTTGAGGGTACAAACTTACAGCACTGGTTCGAAAGATTGG 650
216 eSerGluSerPheProProValSerLeuHisTyrSerTyrAspLeuArg 233
651 TCAAGACATTTTGGCCCCAAATATCTCCGGGGTTTTCTGTATGATCCACT 700
233 erLeuValSerGluLeuPheLysArgAsnProArgAspArgProSerVal 249
701 CCTGTATATCTAGCTCTTTCAAGATATCTCTCCGAGACCCGACATTCATA 750
250 AsnSerIleLeuGluLysGluSoyPheIleAlaLysArgIleGluCysPhe 266
751 AATTCATTTTGAAGAGCCCTTTTACAGAACTATATCCSAAATATT 800
266 uSerProGln.....LeuIleAlaGluArg 275
801 GACTCTGAGGTAAGTTTGAAGTACACGTGGTGGATTGGSCAAAGATT 850
275 heCysLeuLysrhrPheSerLys.....PheGlySerGlnProIle 288
851 TGGGTTGCAGTCTTGACACAGCTGTCTGGTTTAAAGTGTGACACATG 900
289 ...ProAla.....LysArgProAlaSerGluIleAsnSe 299
901 CACCCGACCTGATATGAGTTTTTAAAAACAATCCSCCTTTGTTCCTCA 950
299 rIleSerValMetProAlaGlnLysIlePheLysProAlaIleAlaLysTyrG 316
951 GATAAATAATGATAGAAAGACCCAAATATGCTGTCTGTGGAATATTAG 1000
316 LysIleProLeuAlaTyrLysLysTyrGluArgLysLysLeuHisGluCys 332
1001 ATTATATTATATGCTCAACTGATATGTCGAGAGAGAGAGCCSAC.... 1044
333 LysArgProLeuGlnLysHisLysGluAlaHisGlnPheProGluLysArgVal 349
1045 AAACCA.....AGTATATACCCCTATCTSCAAGAA.... 1074
349 LysAsnThrGluGluAlaArgArgLysIleSerGluGluAlaAlaArgLysA 366
1075 .AATACTGAGATTGAG..... 1089
366 rGArgLeuGluIlePheIleGluLysGluLysLysGlnLysAspGlnIleIle 382
1090 .....GATTACGCTGAGAAACGAGAGCATGTGCATCCSCA.... 1125
383 SerLeuMetLysAlaGluIleMetLysrArgGlnGlnLysGluArgLeuG 399
1116 AGTCATATGGCCGTGTGATGACTCTCAGAGAAATAATTTGAAGTCAACATA 1175
399 uArgIleAsnArgAlaArgGluGlnGlnTyrPArgAsnValLeuSerAlaG 416
1176 TAAGTTGAAGAGTGGAGAACCAATTGGGCTTCTGT..... 1209
416 LysLysSerGluGluValLysAlaProPheLeuGluSerGlyLysIleThrIle 432
1209 ..... 1209

```

433	AlAPserSerPheSerSerArgGlyGlnTyrGlnHisTyrHisAla	449
1210	...CCATCTTCCGCGCCAAAT	1230
449	ePheAspGlnMetGlnGlnGlnArgAlaGlnAspAsnGlnAlaTyrTrp	466
1231	..TTCACCCAGACAGACACTAAGAAAGTAAATGAGACAGACGCTGAGATTC	1279
466	ysArgGlnTleTyrGlnArgGlyLeuProGluArgGlnLysGlyGlnLeu	482
1280	AGGAG.....CTCCCATTTTAAAGAAACGAAATGAG	1311
483	AlaValGlnArgAlaLysGlnValGlnGlnPheLeuGlnArgLysArgG	499
1312	GAACAGACAAATATTTGACGACGTTAAGAGAA.....	1341
499	uAlaMetGlnAsnLysAlaArgAlaGlnGlnHisMetValTyrLeuAla	516
1341	1341
516	rGluArgGlnLleArgGlyLeuGlnAsnPheAsnGlnArgGlnGlnLle	532
1342ATACCCCAACGTCACCAACATGACATGATGAAGAAATTTAGA	1380
533	AlaLysLeuArgGlyGlyLysGlnAlaAsnHisSerGlnGlyGlnG	549
1381	AGAGAGATGGCGAGAACACAGAGCAAT.....	1410
549	uGlySerGlnGlnAlaAspMetArgArgLysLysLleGlnSerLeuLys	566
1411GAAAAGACCTTGAAACAAATGAGCTTCAGAACCAACAG	1449
566	lHisAlaAsnAlaArgAlaAlaValLeuLysGlnGlnLeuGlnArgLys	582
1450GAAAGTAA	1458
583	ArgLysGlnLAlaTyrGlnArgGlnLysValTTrpGlnGlnHisLeuVal	599
1459	AAATCCAGAA.....CGAAATTAATAAGC	1481
599	lAlaLysGlyAlaLysSerSerAspValSerProLeuLeuGlnGlnHisG	616
1482	TAAAGAGGGGTAAATTT.....	1500
616	lunThrGlySerProSerLysGlnGlnMetArgSerValIleSerVal	632
1500	1500
633	ThrSerAlaLeuGlyGlnValAlaAspSerSerLeuTrpAspThr	649
1501CAATTAATTTAGCAAAATGATATTTGAT.....	1530
649	gGluTrpSerGlnGlnMetGlnLysThrAsnAlaIleSerSerLys	666
1530	1530
666	rGlnLleLeuArgArgLeuAsnGlnAsnLeuLysAlaGlnGlnAspGln	682
1531GAAACATCTCCAGAGAGAGAGCA	1557
683	LysGlyMetClnAsnLeuSerSphTrpPheGln.....IleAs	695
1558	ATGCAATATACCAATATAAACTTTGACCTTTGAGATGCGCATGAAGTTTAA	1607
695	nValHisGlnAspAlaLysGlnHis.....G	704
1608	GGAATATGAATGTGTAAAGAGCATGAGCATTTATACAGACAAAGCATTTG	1657
704	lLys.....GlnLysSerValSerSer.....	711
1658	AAAAACTTCACTCCCAAGAGAGGTTTTCACACGACACTGTAGCTGCT	1707
712AspArgLysLysTrpGlnAlaGlyGln.....GlnLeuValIleLeu	725

[illegible]

Identify to protein.

FEATURES
source

Location/Qualifiers
1..2134
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:29949 IMAGE:5000918"
/tissue_type="Placenta, chorioncarcinoma"
/clone_lib="NH.MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
272..1624
/product="unknown (protein for MGC:29949)"
/protein_id="AAH19916.1"
/db_xref="GI:18043743"

CDS

/translation="MKHPVAFKSEFAEHLIYMEYCDGDLMOKITOOKKLEP
EDILMFTOMKLVNHLHKKVLRDLSKNIFLTONKVKLDFGSARLISNPVAF
ACTYVGTPTYPEPEIWMENLYNNKSDIWSLDCILYELCTLHPRQANSWNLKVCQ
GCISLPISHSYELOPLVKOMKRNPSHRPSATLSGIVARLVORCLPEITMEG
EVLLEIKSKKNTPRKTPSRIRIALGNEASVQEEEDPKGSHDLSEINENIVE
SALRRVNRREKGNKSVHLRKASSPNLHRROKVPNTALALENASILTSILAEED
RGSVIVKSKNTKRWOMLETPDLINLIKLNADLSFQYTYIRPSEGLKGPLSE
ETEASDSYDGHDSVILDPERLEPLDEDEDYDFEEDNDPMVSELKRRAGMOGLCDR
"

BASE COUNT 678 a 416 c 494 g 546 t
ORIGIN

alignment_scores:
Quality: 882.00 Length: 412
Ratio: 3.031 Gaps: 6
Percent Similarity: 70.631 Percent Identity: 42.476

alignment block:
US-09-783-320-4 x BC019916 ..

Align seg 1/1 to: BC019916 from: 1 to: 2134

1 MetGluLysTyrValArgLeuGlnLysIleGlyGluGlySerPheGly 17
104 ATGGATGACATACATGCTCTGAGATGATGGGAGGGCTCTCGGCAC 153
17 sAlaIleuValLysSerThrGluAspGlyArgGlnTyrValIleLysG 34
154 AGCTCTTTGGTTCAGCATGAACAGTAAATCATGTTGCCATGAAAG 203
34 IuIleAsnIleSerArgMetSerSerLysGluArgGluGluSerArg 50
204 AAATAGAGGCTTCCCAAG...TCTTCTCTAATACACAGAAATCTAGAG 250
51 GluValAlaValIleuAlaAsnMetLysHisProAsnIleValGlnTyr 67
251 GAGGCTGTTCTTTAGCCAAATGAACACCCATAATTTGTGGCTCA 300
67 gGluSerPheGluGluAsnGlySerLeuTyrIleValMetAspTyrCys 84
301 AGAATCATTTGACGTGAAGACACTGTATATTGTGATGATGACTGTC 350
84 LuGlyGlyAspLeuPheLysArgIleAsnAlaGlnLysGlyValLeuPhe 100
351 ATGGAGGGGATCTAATGCAAAAGATTAACACGCAAGAAAGAAAGTAT 400
101 GlnGluAspGlnIleLeuAspTyrPheValGlnIleCysLeuAlaLeu 117
401 CCTGAAGACATGATCTAATGTTTAAACCAAAATGTCCTTGAGATA 450
117 sHisValHisAspArgLysIleLeuHisArgAspIleLysSerGlnAsn 134
451 TCACATTCACAAGAAACCTGTGTACACAGATATCAAGTCCACAAGATA 500
134 LePheLeuThrLysAspGlyThrValGlnLeuGlyAspPheGlyIle 150
501 TCTTCTCATCTCAGAAATGAAATGAAATTTGGAGACTTTGGATCTGCC 550

151 ArgValLeuAsnSerThrValGluLeuAlaArgThrCysIleGlyThrPr 167
551 CGTCTCTCTCCAAATCCAGATGCATTTGCTGTACTTACTGATGGCAATCC 600
167 oTyrTyrLeuSerProGluIleCysGluAsnLysProTyrAsnAsnLys 184
601 TTATTATGTCCTCCAGAAATTTGGAAACCTGCTTATACATATAAA 650
184 eAspIleThrPalaLeuGlyCysValLeuTyrGluLeuCysThrLeuLys 200
651 GTGACATCTGCTCTGCTGGTGCATCTGTATGACTGTATACCTTAAAG 700
201 HisAlaPheGluAlaGlySerMetLysAsnLeuValIleLysIleIle 217
701 CATCATTTAGGCAATAGTGTGAAATCTTATCTCAAGATGATGCA 750
217 rGlySerPheProProValSerLeuHisTyrSerTyrAspLeuArgSerL 234
751 AGGCTGCATCAGTCCACCTGCCCTCATTCATTCATGAACTTCAGTCC 800
234 euValSerGlnLeuPheLysArgAsnProArgAspArgProSerValAsn 250
801 TAGTCAAGCAGATGTTAAAGGAATCCTCACATCGCCCTCGCTACA 850
251 SerIleLeuGluLysGlyPheIleAlaLysArgIleGluLysPheLeu 267
851 ACCGTTCTCTCGAGCATCGTAGCTCGGCTTGCCAGAAAGTCTTACC 900
267 rProGlnLeuIleAlaGluGluPheCysLeuLysThrPheSerLysPhe 284
901 CCCGAGATCATCATGAA..... 919
284 LysSerGlnProIleProAlaLysArgProAlaSerClyGlnAsnSerIle 300
919 919
301 SerValMetProAlaGlnLysIleThrLysProAlaAlaLysTyrGlyI 317
919 919
317 eProLeuAlaLysLysLysTyrGlyAspLysLysLeuHisGluLysP 334
920TATGTGAGGAGAGATTTAGAAAGAAATTA. 949
334 roLeuGlnLysHisLysGlnAlaHisGlnThrProGluLysArgValAsn 350
950AATTCAGACCATACACACCAAGAAAGAAACAAAC 985
351 ThrGlyGluGluArgArgLysIleSerGluGluAlaAlaArgLysArg 367
986 CCCAGCAGATTCAGATAGCTTTGGAAATGAACAGACAGACAGTCCAA. 1033
367 gLeuGluPheIleGluLysGluLysLysGlnLysAspGlnIleIleSerL 384
1034GAGGAAGAACAGATGAAAGAGGTATACCTATGATT 1070
384 euMetLysAlaGluGlnMetLysArgGlnGluLysGluArgLeuGluArg 400
1071 TGGAAACCATTAATGAA.....AATTTAGTTGAAGTCAATTTGAGAGA 1114
401 IleAsnArgAlaArgGluGlnGlyTyrPargAsnVal 412
1115 GTAAACAGA...GAGAGAAAGGTAATTAAGTCACTC 1147

seq_name: gb-pr:HUMSTRK2A

seq_documentation_block:

LOCUS HUMSTRK2A 3698 bp mRNA linear PRI 09-AUG-1994
DEFINITION Human protein serine/threonine kinase stk2 mRNA, complete cds.
ACCESSION L20321
VERSION L20321.1 GI:348244
KEYWORDS protein serine/threonine kinase; serine/threonine kinase.


```

14002 ...ATPAGAAAAAAGAAATTGGAAGCCCGTGTGCGGAGGACGCAAAACT 13956
397 gLeuGluArg.....IleasnArgAlaArgGlu 407
13955 ACTGTGAAAAGCGAGCAAGCGCAGCAGCTGTGGAGAAACAAAGGAGAG 13906
407 InGlyTrpArgAsnValLeuSerAlaGlyGlySerGlyGluValTyrAla 423
13905 AGCAATCG..... 13898
424 ProPheLeuGlySerGlyGlyThrIleAlaProSerSerPheSerSerAla 440
13897 .....GAGCG 13893
440 gGlyGlnIleArgIleHisAlaIlePheAspGlnMetGlnGlnAla 457
13892 CAACATGAGGAGCAT.....GCAGAGGAAGTTAGGCGAGGCC 13855
457 rGAlaGluAspAsnGluAlaLeuTrpLysArgGluIle..... 469
13854 AAGGAGAGGAGGCTGTGCACGTGACAAACAGAAACAAACGCTCTAAC 13805
470 .....TyrGlyArgGlyLeuProGluArgGlnLysGlu 480
13804 GATAGACCGGTAAATCGATACGGTGAATGCT...GCACGCGCGTCTCATTT 13758
480 pGlnLeuAlaValAlaGluArgAlaLysGlnValGlnGluPheLeuGlnArgL 497
13757 GAAATGCCACCTCTCTCCACAGCAGCAGCAGAGTTTACCCTGAGATGC 13708
497 yArgGluAlaMetGlnAsnIleAlaArgAlaArgLysGlnMetValTyr 513
13707 GCCCGCAGGACCCCTGAAATAG..... 13685
514 LeuAlaArgLeuArgGlnIleArgLeuGlnAsnPheAsnGluArgGlnGlu 530
13684 .....CAGCGCTTATGGAAGAGATGAA 13662
530 nIleLys.....AlaL 534
13661 ACTAGACAGACCACTCTCGCTTACGGGCAACCCACAGCGAGATGGGA 13612
534 yLeu.....ArgGlyGlnLysGlnAlaAsnHis...SerGlu 546
13611 AATTACCAACCGAATAGGGAGGAGAAAGCGCTTGAACCTCTGGGCAAT 13562
547 GlyGlnGlnGlySer.....GlnGluAlaAspMetArgArg 558
13561 GGTCAGCAGCGCCACACGATATCAACACTAGACATATGGAAGCAGCCCG 13512
558 glySylSylLeuSerLeuLys.....AlaHisAlaAsnAlaArgAla 573
13511 CAAGGCTTATTTGGATATGCGGCGGAGGACAGAGACAGCGGACGAG 13462
573 lValLeuLysGlu.....GlnLeuGluArgLysArg 583
13461 CACTGCTTATGAGCAGCTCCAAAGCGCTGTGAGTGGAACAAACC... 13415
584 LysGluAlaTyrGluArgGluLysLysValTrpGluGlnHisLeuValAla 600
13414 GAAGCGGACACCCCTAAGAAATCGCAATG.....CCACACATATGAA 13371
600 aLysGlyVal.....LysSerSer 607
13370 AAAGCTGACACACAGCAAGCGCCGCAAGCAATGATACAAACAAAGC 13321
607 sPValSerProPheLeuGlyGlnHisGluTrpGly.....GlySer 620
13320 ATGTTTGCCTTGTCTATGCAAAAGCGCACCGCCAAACAAATGAAGGCCG 13271
621 ProSerLysGlnGlnMetArgSerValIleSerValThrSerAlaLeuLys 637
13270 CCGTGCCACAA.....CTGAGCCGGAACACTCCCGCCAA 13236
637 sGluValGlyVal.....AspSerSerLeuThrAspThrArgGluTrpS 652
13235 TAGGGTTGAGGAGCTACCTCAAGATTAGTGCAGCAATTCCTCAGATAGTG 13186
652 erGlnMetGlnLysThrAsnAsnAlaIleSerSerLysArgGluIle 668
13185 GGACGAG..... 13178
669 LeuArgArgLeuAsnGlnLysAsnLeuLysAlaGlnGluAspGluLysGlyMe 685
13177 .....GTGAAGAGAGACGAGGATTATACGCACT 13149
685 tGlnAsnLeuSerAsp.....ThrPheGluIleAsnValHisGluAspA 700
13148 CCAGACTCTTATAGACAGGCGCTGCACAAAGAGGCGCATCTCGAGAGTG 13099
700 lAlysGlnHisGluLysGluLysSerValSerSerAspArgLysLysTrp 716
13098 ATGAAGACTTCATATGATGAAGCGCTGTGACTGTGATATATACAGTTT 13049
717 GluAlaGlyGlyGln...LeuValIleProLeuAspGluLeuThrLeuAs 732
13048 GTGCTGATGACGACACTCGCTCCGAAAGTGCAGCACCA..... 13004
732 pThrSerPheSerThrThrGluArgHisThrValGlyGluValIleLysL 749
13003 .....GATCCATATGATGACCGTATC...GAGTCTCTTCGCC 12970
749 euGlyProAsnGlySerProArgArgAlaTrpGlyLysSerProThrAsp 765
12969 TTTTCTTGAAGCGAA.....CTGGCAGATCC..... 12941
766 SerValLeuLysIleLeuGlyGluAlaGluLeuGlnLeu..... 778
12940 .....AAACTTTGGA...GCTTACCGCGCATGATGATATATTAG 12903
779 .....GlnThrGluLeuLeuGlnLysAsnThrThrIleArgSerGluLys 793
12902 TGCAGAGCAGCAGCAATGCAATGCAAGCT.....GTTCCGAGTGAACCTTC 12859
793 erProGlnGlyGluLysTyrLysProLeuIle 803
12858 CAGAAGCCACCAATAAATACATCCCAATCATTT 12827

```

THIS PAGE BLANK (uspro)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 07:03:33 ; Search time 54.14 Seconds
(without alignments)

547,704 Million cell updates/sec

Title: US-09-783-320-4

Sequence: 1 MEKYVRLQKIGSGFGKAL.....YAKILHLMADGAYQEDNDE 1214

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :

Issued_Patents-AA:*
1: /cgn2_6/prodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/prodata/1/iaa/Backfilltest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1734	27.8	345	US-09-173-581-6	Sequence 6, Appl 1
2	1734	27.8	345	US-09-420-915-6	Sequence 2, Appl 1
3	599.5	7.8	979	US-08-870-529-2	Sequence 2, Appl 1
4	489.5	7.8	1360	US-09-393-569-2	Sequence 2, Appl 1
5	485.5	7.8	302	US-09-221-235-2	Sequence 2, Appl 1
6	485.5	7.8	302	US-09-221-928-2	Sequence 2, Appl 1
7	485.5	7.8	302	US-09-221-527-2	Sequence 2, Appl 1
8	485.5	7.8	302	US-09-221-236-2	Sequence 2, Appl 1
9	485.5	7.8	302	US-09-221-245-2	Sequence 2, Appl 1
10	485.5	7.8	302	US-09-163-115-2	Sequence 2, Appl 1
11	485.5	7.8	302	US-09-221-528-2	Sequence 2, Appl 1
12	485.5	7.8	302	US-09-593-553-2	Sequence 2, Appl 1
13	485.5	7.8	302	US-09-221-237-2	Sequence 2, Appl 1
14	485.5	7.8	302	US-09-112-709-1	Sequence 2, Appl 1
15	485.5	7.8	302	US-09-111-444-1	Sequence 2, Appl 1
16	451	7.2	233	US-09-541-228-1	Sequence 1, Appl 1
17	451	7.2	233	US-08-685-871-1	Sequence 1, Appl 1
18	445	7.0	1354	US-09-049-671-1	Sequence 1, Appl 1
19	436	7.0	293	US-09-060-110-4	Sequence 4, Appl 1
20	436	7.0	293	US-09-060-110-4	Sequence 4, Appl 1
21	432.5	6.9	993	US-09-031-563-5	Sequence 5, Appl 1
22	430	6.9	647	US-09-031-563-5	Sequence 5, Appl 1
23	430	6.9	648	US-09-031-563-5	Sequence 5, Appl 1
24	430	6.9	1315	US-09-293-505-10	Sequence 25, Appl 1
25	430	6.9	1315	US-09-293-505-10	Sequence 25, Appl 1
26	427	6.8	1315	US-08-685-576-4	Sequence 4, Appl 1
27	426	6.8	1386	US-08-685-576-4	Sequence 4, Appl 1

28	425	6.8	1388	US-08-685-576-1	Sequence 1, Appl 1
29	412	6.6	260	US-09-031-563-24	Sequence 24, Appl 1
30	407.5	6.5	925	US-08-252-995D-4	Sequence 4, Appl 1
31	407.5	6.5	925	US-08-834-108-4	Sequence 4, Appl 1
32	405	6.5	1001	US-09-060-410-2	Sequence 62, Appl 1
33	397	6.4	900	US-08-630-822A-62	Sequence 62, Appl 1
34	397	6.4	900	US-09-005-069-62	Sequence 62, Appl 1
35	384	6.2	239	US-08-712-709-7	Sequence 7, Appl 1
36	384	6.2	239	US-09-049-671-3	Sequence 3, Appl 1
37	384	6.2	239	US-09-295-068-3	Sequence 3, Appl 1
38	384	6.2	239	US-09-111-444-7	Sequence 7, Appl 1
39	384	6.2	239	US-09-541-228-7	Sequence 7, Appl 1
40	384	6.2	729	US-08-677-298-2	Sequence 2, Appl 1
41	381	6.1	416	US-08-252-995D-2	Sequence 2, Appl 1
42	381	6.1	416	US-08-834-108-2	Sequence 2, Appl 1
43	381	6.1	464	US-08-252-995D-6	Sequence 6, Appl 1
44	381	6.1	464	US-08-834-108-6	Sequence 6, Appl 1
45	380.5	6.1	685	US-08-878-989-1	Sequence 1, Appl 1

ALIGNMENTS

RESULT 1
US-09-173-581-6
Sequence 6, Application US/09173581A
Patent No. 6013455
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Gorgone, Gina
APPLICANT: Azimzai, Yalda
APPLICANT: Lu, Aina
TITLE OR INVENTION: Protein Kinase Homologs
FILE REFERENCE: PR-0614 US
CURRENT APPLICATION NUMBER: US/09/173,581A
CURRENT FILING DATE: 1998-10-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PERL Program
SEQ ID NO 6
LENGTH: 345
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: -
OTHER INFORMATION: 1567782
US-09-173-581-6

Query Match 27.8%; Score 1734; DB 3; Length 345;
Best Local Similarity 90.5%; Pred. No. 1.3e-89;
Matches 341; Conservative 3; Mismatches 1; Indels 32; Gaps 1;

QY 1 MEKYVRLQKIGSGFGKALVKSTEDGRQYVIREINISRMSSKERESEREYAVLANKH 60
DB 1 MEKYVRLQKIGSGFGKALVKSTEDGRQYVIREINISRMSSKERESEREYAVLANKH 60
QY 61 PNIVYRSEFENGSLTYIMDYCEGDLFFKRNQKGVLFQEDQIDMFWQICLALKHVH 120
DB 61 PNIVYRSEFENGSLTYIMDYCEGDLFFKRNQKGVLFQEDQIDMFWQICLALKHVH 88
QY 121 DKTIHLROIKSONIFLTGDTGVOGDFGIARVUNSVTELARTICIGPYVLSPEICENRY 180
DB 89 DKTIHLROIKSONIFLTGDTGVOGDFGIARVUNSVTELARTICIGPYVLSPEICENRY 148
QY 181 NKSQDIWALGCVLYELCTLKRAFEAGSMKNLYLKIIISGSPVSLHYSVDLSLSOLEK 240
DB 149 NKSQDIWALGCVLYELCTLKRAFEAGSMKNLYLKIIISGSPVSLHYSVDLSLSOLEK 208
QY 241 RNRDRPSVNSILDEKGFIAKRIEKLFLSLQILAEFCLTKTSKSGSDIPAKRPAASQNSI 300

Wed May 15 14:13:55 2002

us-09-783-320-4.ra1

Page 2

```
Db 209 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEEFLCTPSKGSOPIPAKRPASGNSI 268
OY 301 SVMPAOKITTKPAAYGIPLAYKKYGDKKLHEKKRPLOKHKQAHOTPEKRVNTGEERKISE 360
Db 269 SVMPAOKITTKPAAYGIPLAYKKYGDKKLHEKKRPLOKHKQAHOTPEKRVNTGEERKISE 328
OY 361 EAARRRLEFTEKEKKO 377
Db 329 EAARRRLEFTEKDER 345

RESULT 2
US-09-420-915-6
Sequence 6, Application US/09420915
Patent No. 626497
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Gorgone, Gina
APPLICANT: Azimzal, Yalda
APPLICANT: Lu, Alina
TITLE OF INVENTION: Protein Kinase Homologs
FILE REFERENCE: PE-0614 US
CURRENT APPLICATION NUMBER: US/09/420,915
EARLIER FILING DATE: 1999-10-20
EARLIER APPLICATION NUMBER: US 09/173,581
NUMBER OF SEQ. ID NOS: 18
SOFTWARE: PERL Program
SEQ. ID NO. 6
LENGTH: 345
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: 1567782
US-09-420-915-6
```

```
Query Match 27.8%; Score 1734; DB 4; Length 345;
Best Local Similarity 90.5%; Pred. No. 1.3e-89;
Matches 341; Conservative 3; Mismatches 1; Indels 32; Gaps 1;

OY 1 MEKYVRLQKIGESFGKAILVSTEDGROYVYKEINISRMSSKEREESRREVAVALANMKH 60
Db 1 MEKYVRLQKIGESFGKAILVSTEDGROYVYKEINISRMSSKEREESRREVAVALANMKH 60
OY 61 PIVVOYRESFEENSLVYMDYCEGDLFRINAKGVLPQEDDILDMFVOICLAKKHVH 120
Db 61 PIVVOYRESFE-----GIIDMVFVOICLAKKHVH 88
OY 121 DRKILHRKISONIFLRKDGTVOLGDEGIAVNSTVELARTGIPPYLSPEICENKRY 180
Db 89 DRKILHRKISONIFLRKDGTVOLGDEGIAVNSTVELARTGIPPYLSPEICENKRY 148
OY 181 NKSSTVALGCVLYELCTLKAHFAFGSMKMLVLIISGSPVSLHSTDLNSLVQLFK 240
Db 149 NKSSTVALGCVLYELCTLKAHFAFGSMKMLVLIISGSPVSLHSTDLNSLVQLFK 208
OY 241 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEEFLCTPSKGSOPIPAKRPASGNSI 300
Db 209 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEEFLCTPSKGSOPIPAKRPASGNSI 268
OY 301 SVMPAOKITTKPAAYGIPLAYKKYGDKKLHEKKRPLOKHKQAHOTPEKRVNTGEERKISE 360
Db 269 SVMPAOKITTKPAAYGIPLAYKKYGDKKLHEKKRPLOKHKQAHOTPEKRVNTGEERKISE 328
OY 361 EAARRRLEFTEKEKKO 377
Db 329 EAARRRLEFTEKDER 345
```

```
Db 329 EAARRRLEFTEKDER 345

RESULT 3
US-08-870-529-2
Sequence 2, Application us/08870529
Patent No. 6080557
GENERAL INFORMATION:
APPLICANT: Sims, John E.
APPLICANT: Virca, G. Duke
APPLICANT: Baird, Timothy A.
APPLICANT: Anderson, Ditt M.
TITLE OF INVENTION: IL-1/TNF-(ACTIVATED KINASE (ITAK)
TITLE OF INVENTION: AND METHODS OF MAKING AND USING THE SAME
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/870,529
FILING DATE: 06-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 418
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 979 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-870-529-2
```

```
Query Match 9.6%; Score 599.5; DB 3; Length 979;
Best Local Similarity 37.5%; Pred. No. 7.4e-26;
Matches 124; Conservative 71; Mismatches 113; Indels 23; Gaps 5;

OY 4 YVRLQKIGESFGKAILVSTEDGROYVYKEINISRMSSKEREESRREVAVALANMKHPT 63
Db 52 YVRLQKIGESFGKAILVSTEDGROYVYKEINISRMSSKEREESRREVAVALANMKHPT 111
OY 64 VOYRESFEENSLVYMDYCEGDLFRINAKGVLPQEDDILDMFVOICLAKKHVH 123
Db 112 VOYRESFEENSLVYMDYCEGDLFRINAKGVLPQEDDILDMFVOICLAKKHVH 171
OY 124 ILHRDRIKSONIFLRKDGTVOLGDEGIAVNSTVELARTGIPPYLSPEICENKRY 183
Db 172 ILHRDRIKSONIFLRKDGTVOLGDEGIAVNSTVELARTGIPPYLSPEICENKRY 231
OY 184 SDIWMALGCVLYELCTLKAHFAFGSMKMLVLIISGSPVSLHSTDLNSLVQLFK 241
Db 232 SDIWMALGCVLYELCTLKAHFAFGSMKMLVLIISGSPVSLHSTDLNSLVQLFK 291
OY 242 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEEFLCTPSKGSOPIPAKRPASGNSI 301
Db 292 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEEFLCTPSKGSOPIPAKRPASGNSI 355
OY 302 SVMPAOKITTKPAAYGIPLAYKKYGDKKLHEKKRPLOKHKQAHOTPEKRVNTGEERKISE 360
Db 355 SVMPAOKITTKPAAYGIPLAYKKYGDKKLHEKKRPLOKHKQAHOTPEKRVNTGEERKISE 413
```


Wed May 15 14:13:55 2002

us-09-783-320-4.rai

Page 4

Matches 102; Conservative 48; Mismatches 101; Indels 11; Gaps 6;

```

Oy 8 QGGESPEFKALVSTEDGKQVYIKRINI--SMSSKEEESRRRVAIVAMKHPNIVQY 66
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 38 KTIIGRGQFESEYVRACCLIDGVPALKKVQIFDLMQAKKARADIKRIDLLKQLNHPNIVKY 97
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Oy 67 RESEFENSGLTYVMYQCGSGD---FKRIAAQGVLFQEDQILMFQDICALHYNDK 123
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 98 YASFIDNENLNLVLELDAGSLSMKIHFKQKR--LIPRIYWKKFVLDCALEHMSKR 156
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Oy 124 ILHRLDKSNPIFLPDGVOVLGDFGRGLARVLANSVLELAFRCITPRYISPELCEKPRYNNK 183
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 157 VMHRDKPRANVTITATGVKLGADGGLDRFFSTSTTAHSLVQPIYMSPEKTIHENGYNFK 216
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Oy 184 SDIWMALGVLYELCTLKIAEPAGSMK--NLVLIITIGSPRY--SLATSYDILRLSVLSOLF 240
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 SDIWMALGVLYELCTLKIAEPAGSMK--NLVLIITIGSPRY--SLATSYDILRLSVLSOLF 276
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Oy 241 RNRDRSPVNSVITLEKGTAKRI 262
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 277 PDQKRRPDVLYVD--VAKRM 295
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 6
US-09-221-928-2
: Sequence 2, Application US/09221928
: Patent No. 6121030
: GENERAL INFORMATION:
: APPLICANT: Acton, Susan
: TITLE OF INVENTION: NOVEL CSAK-1 NOCLEIC ACID MOLECULES AND USES THEREFOR
: FILE REFERENCE: MN-030
: CURRENT APPLICATION NUMBER: US/09/221,928
: CURRENT FILING DATE: 1998-12-28
: EARLIER APPLICATION NUMBER: 09/163,115
: EARLIER FILING DATE:
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 302
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-221-928-2

```

Query Match	7.8%	Score 485.5;	DB 3;	Length 302;
Best Local Similarity	38.9%	Pred. No. 4e-20;		
Matches 102;	Conservative 48;	Mismatches 101;	Indels 11;	Gaps 6

Qy	8	OKIGESSFKGKALVSTEDGQYVIRKINI--SKASSKEKEBESREYAVLANKKHPNIVQY	66
Dh	38	KKIGRGQGESEVYRAACCLIDGPPVAVLKQVLFIDLDKKKAKADCIKEDLDLKQINHPNIYKY	97
Qy	67	RESFEENGSLTYVMDYICGGGL--FRIANQGVLFQEDQILDMFVOICLAKVHDHK	123
Dh	98	YASFIEDNELNVLLEDAAGLSMNIHFKQKR--LIPRTYWKYFVQDCSLEHMSRR	156
Qy	124	ILHRDIKSONIFLYTKDGTVOVGDFGIARVINSTVELARTCIQTPYUUSPEICENKPYNNK	183
Dh	157	VMRHRIKPPANVPIITATGVVYKLGDDGLGRFSSSTTAASHLVQTPYUUSPERIHENGNNFK	216
Qy	184	SDIWAIGCVLYELCTLKHAFEAGSMK--NLVTKIIGSGPPY--SLHYSYDRLSLVQLFK	240
Dh	217	SDIWSIGCLLYEMAAALQSPFGDMKNMYSLCKKIEQCDYPLPSDHSYSEELROLVNMCIN	276
Qy	241	RNPDRPPVNSILKGFATKRI	262
Dh	277	PDEKRPDVTYYD--VAKKM	295

RESULT 7
US-09-221-527-2
; Sequence 2, Application US/09221527
Patent No. 6146832

```

GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MN-050
CURRENT APPLICATION NUMBER: US/09/221,527
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 302
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-527-2

```

Query Match	7.88	Score 485.5	DB 4	Length 302
Best Local Similarity	38.9%	Pred. No. 46-20		
Matches 102	Conservative 48	Mismatches 101	Indels 11	Gaps
Qy	8	QKIGESSEKALIVSTEDGKQYIKELNI-SKMSKEEESREYAVLANKRPNIYQY	66	
	
Db	38	KKIGRQSESEYVRACCLIDGQPVALKQQLFDDMDAAADCKIEDLKQNHNPNIKI	97	
Qy	67	RESFEENGSLIYVMDYCEGQD---FRIANQGVLFQEDQILDMFVQIALCAKHVDRK	123	
	
Db	98	YASFIEDNELNIVLELADQAGDLSMTIKHFQKQR-LIPEIYWKTFVQCSLSEHMSRK	156	
Qy	124	ILHRDIKSONIFLFTKDGTVQJGDFGIARVLNSTVELARTICIGTPYLSPEICEKPPYNNK	183	
	
Db	157	VHMHRDIKPNPNVITATGAVKVLGDLGJLGRFESSKTTAAHSLVCGTPYMSPERIHENGNIKF	216	
Qy	184	SDIWAIGCVLYELCLTKHAFAFGSMK--NLVLKIISSGFPPV-SLHYSYDLRSLSOLEFK	240	
	
Db	217	SDIWSIGCLLYEMALQSPFGDKMNLVSLCKRIEODCVPLPSDHYSEBELROLVNMCIN	276	
Qy	241	RNRDRPPSVNSILKGFATKRI 262		
	
Db	277	PDPKRPDYTVYVD--VAKRM 295		

```

RESULT      8
US-09-221-236-2
: Sequence 2, Application US/09221236
: Patent No. 6146841
GENERAL INFORMATION:
APPLICANT: ACTION, SUSAN
TITLE OF INVENTION: NOVEL CSAFK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-030
CURRENT APPLICATION NUMBER: US/09/221,236
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 302
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-236-2

```

	Query Match	7.8%	Score 485.5;	DB 4;	Length 302;
	Best Local Similarity	38.9%	Pcd No. 4e-20;		
Matches	102; Conservative	48;	Mismatches 101;	Indels 11;	Gaps 6.
OY	8 OKIEGSFGALIVKSTEDGRQVIREIKI--SRMSKRRESREVVVLNMKNHPNV	IYY	66		
Dd	38 KRIIRGFSESVRAACILCGVPALKRVQIPDIIMARKRABICITIEDILLKOLNHNPIVK	I	97		
OY	67 RESFENGSLIIYMDCCEGGD---FKRLNAKGVLFOEQDILMFPOIALAKLHNDKR	I	123		

```

Db      98 YASFIEDNELNIVLELADAGDLSRMKHKOKR-LIPERTVWKYVQLCSALEHMSRR 156
;
Oy      124 ILHRDIKSONIFLTKDGTVOQLDGFIAVLNSTVELARTICIGPYVLSPEICENKPYNNK 183
;
Db      157 VMHRDIKIPANVFITATGVVKKLDLGLGRFSSKTTAAHSLVGPYYMSPERIHENGYNK 216
;
Oy      184 SDIWAIGCVLYELCTLKAHFEAGSMK--NLVLTIGSGFPPV-SLHYSYDLRSIVSOLFK 240
;
Db      217 SDIWSIGCLLYEMALOSPFYGDKNMNLVSLCKRIEODCDYPLPSDHSSELRQVLMCIN 276
;
Oy      241 RNPDRPSVNSILEKGFIAKRI 262
;
Db      277 PDPEKRPDYTYVD--VAKRM 295
;
;
RESULT 9
US-09-221-416-2
; Sequence 2, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-2

Query Match      7.8%; Score 485.5; DB 4; Length 302;
Best Local Similarity 38.9%; Pred. No. 4e-20;
Matches 102; Conservative 48; Mismatches 101; Indels 11; Gaps 6;

Oy      8 OKIGSGFGKAILVSTEDGRQVYKEINT-SRMSKERESESRREAVLANMKHPNIVQY 66
;
Db      38 KTIIGQGFSEVYRAACLDGVPVAKKVOIFDLMDAKARADCIKEIDLKOLNHPNVIKY 97
;
Oy      67 RESFENGSLYIVMDYCEGGD---FKRINAKGVLFQEDQILDMFVQICLAKVHDKR 123
;
Db      98 YASFIEDNELNIVLELADAGDLSRMKHKOKR-LIPERTVWKYVQLCSALEHMSRR 156
;
Oy      124 ILHRDIKSONIFLTKDGTVOQLDGFIAVLNSTVELARTICIGPYVLSPEICENKPYNNK 183
;
Db      157 VMHRDIKIPANVFITATGVVKKLDLGLGRFSSKTTAAHSLVGPYYMSPERIHENGYNK 216
;
Oy      184 SDIWAIGCVLYELCTLKAHFEAGSMK--NLVLTIGSGFPPV-SLHYSYDLRSIVSOLFK 240
;
Db      217 SDIWSIGCLLYEMALOSPFYGDKNMNLVSLCKRIEODCDYPLPSDHSSELRQVLMCIN 276
;
Oy      241 RNPDRPSVNSILEKGFIAKRI 262
;
Db      277 PDPEKRPDYTYVD--VAKRM 295
;
;
RESULT 10
US-09-221-245-2
; Sequence 2, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; EARLIER FILING DATE: 1998-09-29

```

```

; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-245-2

Query Match      7.8%; Score 485.5; DB 4; Length 302;
Best Local Similarity 38.9%; Pred. No. 4e-20;
Matches 102; Conservative 48; Mismatches 101; Indels 11; Gaps 6;

Oy      8 OKIGSGFGKAILVSTEDGRQVYKEINT-SRMSKERESESRREAVLANMKHPNIVQY 66
;
Db      38 KTIIGQGFSEVYRAACLDGVPVAKKVOIFDLMDAKARADCIKEIDLKOLNHPNVIKY 97
;
Oy      67 RESFENGSLYIVMDYCEGGD---FKRINAKGVLFQEDQILDMFVQICLAKVHDKR 123
;
Db      98 YASFIEDNELNIVLELADAGDLSRMKHKOKR-LIPERTVWKYVQLCSALEHMSRR 156
;
Oy      124 ILHRDIKSONIFLTKDGTVOQLDGFIAVLNSTVELARTICIGPYVLSPEICENKPYNNK 183
;
Db      157 VMHRDIKIPANVFITATGVVKKLDLGLGRFSSKTTAAHSLVGPYYMSPERIHENGYNK 216
;
Oy      184 SDIWAIGCVLYELCTLKAHFEAGSMK--NLVLTIGSGFPPV-SLHYSYDLRSIVSOLFK 240
;
Db      217 SDIWSIGCLLYEMALOSPFYGDKNMNLVSLCKRIEODCDYPLPSDHSSELRQVLMCIN 276
;
;
RESULT 11
US-09-163-115-2
; Sequence 2, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-163-115-2

Query Match      7.8%; Score 485.5; DB 4; Length 302;
Best Local Similarity 38.9%; Pred. No. 4e-20;
Matches 102; Conservative 48; Mismatches 101; Indels 11; Gaps 6;

Oy      8 OKIGSGFGKAILVSTEDGRQVYKEINT-SRMSKERESESRREAVLANMKHPNIVQY 66
;
Db      38 KTIIGQGFSEVYRAACLDGVPVAKKVOIFDLMDAKARADCIKEIDLKOLNHPNVIKY 97
;
Oy      67 RESFENGSLYIVMDYCEGGD---FKRINAKGVLFQEDQILDMFVQICLAKVHDKR 123
;
Db      98 YASFIEDNELNIVLELADAGDLSRMKHKOKR-LIPERTVWKYVQLCSALEHMSRR 156
;
Oy      124 ILHRDIKSONIFLTKDGTVOQLDGFIAVLNSTVELARTICIGPYVLSPEICENKPYNNK 183
;
Db      157 VMHRDIKIPANVFITATGVVKKLDLGLGRFSSKTTAAHSLVGPYYMSPERIHENGYNK 216
;
Oy      184 SDIWAIGCVLYELCTLKAHFEAGSMK--NLVLTIGSGFPPV-SLHYSYDLRSIVSOLFK 240
;
Db      217 SDIWSIGCLLYEMALOSPFYGDKNMNLVSLCKRIEODCDYPLPSDHSSELRQVLMCIN 276
;

```



```
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
; US-08-712-709-1
```

```
Query Match 7.28; Score 451; DB 2; Length 233;
Best Local Similarity 38.98; Pred. No. 2.4e-18;
Matches 91; Conservative 48; Mismatches 85; Indels 10; Gaps 5;

QY 40 MSKEREBSREAVLANMKHPNIVOTRESFEENGSLTYMDYCEGDL---FKRINAK 96
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2 MAAKAKODCVKEIGLKQINHPNITKYIDSFIEDNELNIVLELADAGDLPOMIKYFKK 61
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 97 GVLFQEDQILDFVQICLALKVHDRIKLNHDIKSONIFLTKDGVQLGDFGIARYLNST 156
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 62 R-LIPERTVWKYFVQLCSAVEHMSRRVMHRDIPANVFTTATGVVKLGDLGGRFSS 120
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 157 VELAPTCIGTPYVLSPEICENKPYNNKSDIWAIGCVLYELCTLKHAFEAGSMK--NLVIK 214
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 TTAHSLVGTGYPMSPERHENGYNFKSDIMSLGCLLYEMALQSPYGDKNLFSLCOK 180
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 215 IIGSGPPY-SLHYSYDLRSLVSOLFRRNPRDRPSVNSILEKGFIAKRIEKL 267
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 IEQCDYPLPGEHYSEKLELYSMCTCPDPPQRPDYGXVH--VAKQMIHMS 231
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: May 15, 2002, 07:53:47
Job time: 3014 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 07:53:00 ; Search time 262.69 Seconds
(without alignments)

1626.649 Million cell updates/sec

Title: US-09-783-320-4
Perfect score: 6243
Sequence: 1 MEKVRLOKIGESFGKAIL.....YAKILHLVADGAYOEINDE 1214

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main.*
1: /cgn2_6/ptodata/1/paa/PCUS.COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06.COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07.COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US082.COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US083.COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US084.COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US085.COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085.COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086.COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087.COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088.COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US090.COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US092.COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US093.COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US094.COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US095.COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US096.COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US097.COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US098.COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US099.COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US100.COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US101.COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US101.COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101.COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US101.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6243	100.0	1214	21	US-09-783-320-4
2	6237	99.9	1214	18	US-09-488-725A-2356
3	6213	99.5	1242	18	US-09-488-725A-2355
4	5162	82.7	1007	21	US-09-783-320-6
5	5138	82.3	1035	21	US-09-783-320-2
6	4735.5	75.9	966	18	US-09-488-725A-5927
7	4735.5	75.9	966	18	US-09-488-725A-5928

8	3787	60.7	778	26	US-60-299-426-30	Sequence 30, Appl
9	2785	44.6	616	22	US-09-824-583-4	Sequence 4, Appl
10	2785	44.6	616	22	US-09-873-404-4	Sequence 4, Appl
11	1734	27.8	345	22	US-09-870-962-6	Sequence 6, Appl
12	1283	20.6	254	22	US-09-715-427-35	Sequence 35, Appl
13	1283	20.6	254	22	US-09-898-837A-35	Sequence 35, Appl
14	1146	18.4	256	21	US-09-760-446A-1590	Sequence 1590, Ap
15	1060	17.0	640	22	US-09-824-583-2	Sequence 2, Appl
16	1057	16.9	640	22	US-09-873-404-2	Sequence 2, Appl
17	1057	16.9	640	22	US-09-760-446A-1590	Sequence 2, Appl
18	1057	16.9	640	22	US-09-760-446A-1590	Sequence 2, Appl
19	882	14.0	506	1	PCT-US01-04098A-1006	Sequence 1006, Ap
20	875	14.0	527	1	PCT-US01-04098A-1294	Sequence 2924, Ap
21	848	13.6	836	26	US-60-208-965-221	Sequence 221, Ap
22	847.5	13.6	841	19	US-60-213-178-724	Sequence 724, Ap
23	847.5	13.6	841	19	US-60-213-178-724	Sequence 1183, Ap
24	841.5	13.5	269	26	US-60-200-366-127	Sequence 127, Ap
25	841.5	13.5	269	26	US-60-248-798-347	Sequence 347, Ap
26	806.5	12.9	510	26	US-60-106-889-6	Sequence 6, Appl
27	790.5	12.7	459	21	US-09-538-992-1182	Sequence 1182, Ap
28	790.5	12.7	459	21	US-09-771-161A-198	Sequence 198, Ap
29	786	12.6	458	26	US-60-200-366-128	Sequence 128, Ap
30	780.5	12.5	416	21	US-09-731-231-6	Sequence 6, Appl
31	758.5	12.1	264	26	US-60-206-600-158	Sequence 158, Ap
32	756.5	12.1	399	26	US-60-213-178-776	Sequence 776, Ap
33	747	12.0	143	1	PCT-US00-26376-111	Sequence 111, Ap
34	743.5	11.9	260	21	US-09-731-231-5	Sequence 5, Appl
35	733.5	11.7	255	21	US-09-715-427-36	Sequence 36, Appl
36	733.5	11.7	255	21	US-09-898-837A-36	Sequence 36, Appl
37	678.5	10.9	234	26	US-60-248-798-113	Sequence 113, Ap
38	622.5	10.0	698	23	US-09-995-530-15	Sequence 15, Appl
39	621.5	10.0	692	23	US-09-992-481-2	Sequence 2, Appl
40	620.5	9.9	692	23	US-09-910-150-5	Sequence 5, Appl
41	604	9.7	645	21	US-09-731-231-2	Sequence 2, Appl
42	598.5	9.6	645	21	US-09-731-231-2	Sequence 2, Appl
43	598.5	9.6	645	21	US-09-731-231-2	Sequence 2, Appl
44	598.5	9.6	645	21	US-09-731-231-2	Sequence 2, Appl
45	598.5	9.6	645	21	US-09-731-231-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-783-320-4
; Sequence 4, Application US/09783320
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hu, Yi
; APPLICANT: Nepomichy, Boris
; APPLICANT: Turner, C. Alexander Jr
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0137-USA
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: US 60/183,582
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/184,014
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1214
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-783-320-4

Query Match 100.0%; Score 6243; DB 21; Length 1214;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1214; Conservative 0; Indels 0; Gaps 0;

Wed May 15 14:13:57 2002

us-09-783-320-4.1.ram

Page 2

```
QY 1 MEKYVRLQKIGESFGKALIVKSTEDGRQYVIKEINISMSKREESREAVLANMKH 60
Db 1 MEKYVRLQKIGESFGKALIVKSTEDGRQYVIKEINISMSKREESREAVLANMKH 60
QY 61 PNIVQYRESEFENGSLIYMDYCEGGDLFKRINAKGVLFQEDDILDMFVOICLAKH 120
Db 61 PNIVQYRESEFENGSLIYMDYCEGGDLFKRINAKGVLFQEDDILDMFVOICLAKH 120
QY 121 DRKILHRDIXSONIFLTQDGVOLGDFGIARVLNSTVELARTCTGTPYILSPICENKPY 180
Db 121 DRKILHRDIXSONIFLTQDGVOLGDFGIARVLNSTVELARTCTGTPYILSPICENKPY 180
QY 181 NNSSDIHALGCVLYELCTLKHAFEGSMKNLVKTIISGSPFVSLHYSDRLSVQLFK 240
Db 181 NNSSDIHALGCVLYELCTLKHAFEGSMKNLVKTIISGSPFVSLHYSDRLSVQLFK 240
QY 241 RNPRDRPSVNSILKGFIAKRIEKLSPOLIAEEFCLKTFESKFSQPIPAKRPASGNSI 300
Db 241 RNPRDRPSVNSILKGFIAKRIEKLSPOLIAEEFCLKTFESKFSQPIPAKRPASGNSI 300
QY 301 SVMPAQKITRPAKYGIPLAYKKYGDGKLHEKKPLQKHQAHOPEKRVNTGEEERKISE 360
Db 301 SVMPAQKITRPAKYGIPLAYKKYGDGKLHEKKPLQKHQAHOPEKRVNTGEEERKISE 360
QY 361 EAARKRLFEIEKEKQKQDOIIISLMKAEOMKROEKERLERINARAEQGMNVLASGSGE 420
Db 361 EAARKRLFEIEKEKQKQDOIIISLMKAEOMKROEKERLERINARAEQGMNVLASGSGE 420
QY 421 VKAPFLSGGTIAPSSFGSGYEHYALFDDMOOORADENAKKREIYGRGLPERQKG 480
Db 421 VKAPFLSGGTIAPSSFGSGYEHYALFDDMOOORADENAKKREIYGRGLPERQKG 480
QY 481 QLAVERAKOYEELQKREAMONKARABEGHYALRLQRIQONNEGOIKAKLRGSKK 540
Db 481 QLAVERAKOYEELQKREAMONKARABEGHYALRLQRIQONNEGOIKAKLRGSKK 540
QY 541 EAHNSEOEGSEADMRKRIESTIKHANARAALVKEOLEERKRAYREKKVMEHEHYA 600
Db 541 EAHNSEOEGSEADMRKRIESTIKHANARAALVKEOLEERKRAYREKKVMEHEHYA 600
QY 601 KGVKSSVSPPLGQHEHGGSPSQOMRSYISTSALEKGVUSSLTDRRESEEMOKTNN 660
Db 601 KGVKSSVSPPLGQHEHGGSPSQOMRSYISTSALEKGVUSSLTDRRESEEMOKTNN 660
QY 661 AISSKREILRLNENLKAODEKGMONLSDTFEINVEDAKHEKEKSVSSDRKKWEAGX 720
Db 661 AISSKREILRLNENLKAODEKGMONLSDTFEINVEDAKHEKEKSVSSDRKKWEAGX 720
QY 721 QLVIPDLDELTLDTSFSTTERHTVGEVIKLGPNGSPRRAMGKSPRTSVLKLIGEAELQOT 780
Db 721 QLVIPDLDELTLDTSFSTTERHTVGEVIKLGPNGSPRRAMGKSPRTSVLKLIGEAELQOT 780
QY 781 ELLENTTIRSEISPEGEKKYPLITGEEKVQCSISHEINPAIYDSVEIKSPFESASQOM 840
Db 781 ELLENTTIRSEISPEGEKKYPLITGEEKVQCSISHEINPAIYDSVEIKSPFESASQOM 840
QY 841 SLKLEGLLEPDDLEITLQEPGSGTKKDSLPCTITDWAISEKRTKTSQADRTIOEN 900
Db 841 SLKLEGLLEPDDLEITLQEPGSGTKKDSLPCTITDWAISEKRTKTSQADRTIOEN 900
QY 901 EVSEEDGVSSITVDOLSDIHIEPGTNDOSHKCOVDKSVQBPFRKVVHSEHLNLPVOVS 960
Db 901 EVSEEDGVSSITVDOLSDIHIEPGTNDOSHKCOVDKSVQBPFRKVVHSEHLNLPVOVS 960
QY 961 VOCSPSEESFAFRSHSLPRKNKNKNSLLGLSTGLFDANPMKLTCSIPDLSKLFTLM 1020
Db 961 VOCSPSEESFAFRSHSLPRKNKNKNSLLGLSTGLFDANPMKLTCSIPDLSKLFTLM 1020
QY 1021 DVPTVGVRODNLDEIKDENIKESGSDSEIVFEETDTDQEQASMEQLRLRQPEE 1080
Db 1021 DVPTVGVRODNLDEIKDENIKESGSDSEIVFEETDTDQEQASMEQLRLRQPEE 1080
QY 1081 YSEEEESVLKNSDVEPTANGTDVADBDNPPSSSALNBEHWSNDGELIASCECDSDVFN 1140
```

```
Db 1081 YSEEEESVLKNSDVEPTANGTDVADBDNPPSSSALNBEHWSNDGELIASCECDSDVFN 1140
QY 1141 HLEELRLHLEOMGFEKFEVEYERIKALHEDEDENEIECSKIYONILGNHQLYAKILH 1200
Db 1141 HLEELRLHLEOMGFEKFEVEYERIKALHEDEDENEIECSKIYONILGNHQLYAKILH 1200
QY 1201 LVWADGAYOEDNDE 1214
Db 1201 LVWADGAYOEDNDE 1214

RESULT 2
US-09-488-725A-2356
Sequence 2356, Application US/09488725A
GENERAL INFORMATION:
APPLICANT: Hysed Inc
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
FILE REFERENCE: 784FLPCT
CURRENT APPLICATION NUMBER: US/09/488, 725A
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552, 317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US09/598, 042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US09/620, 312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US09/653, 450
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US09/662, 191
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: US09/693, 036
PRIOR FILING DATE: 2000-10-19
PRIOR APPLICATION NUMBER: US09/727, 344
PRIOR FILING DATE: 2000-11-29
NUMBER OF SEQ ID NOS: 7144
SOFTWARE: PL_Fl_genes_D Versions 1.0
SEQ ID NO 2356
LENGTH: 1215
TYPE: PKT
ORGANISM: Homo sapiens
US-09-488-725A-2356

Query Match 99.9% Score 6237; DB 18; Length 1214;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1213; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEKYVRLQKIGESFGKALIVKSTEDGRQYVIKEINISMSKREESREAVLANMKH 60
Db 1 MEKYVRLQKIGESFGKALIVKSTEDGRQYVIKEINISMSKREESREAVLANMKH 60
QY 61 PNIVQYRESEFENGSLIYMDYCEGGDLFKRINAKGVLFQEDDILDMFVOICLAKH 120
Db 61 PNIVQYRESEFENGSLIYMDYCEGGDLFKRINAKGVLFQEDDILDMFVOICLAKH 120
QY 121 DRKILHRDIXSONIFLTQDGVOLGDFGIARVLNSTVELARTCTGTPYILSPICENKPY 180
Db 121 DRKILHRDIXSONIFLTQDGVOLGDFGIARVLNSTVELARTCTGTPYILSPICENKPY 180
QY 181 NNSSDIHALGCVLYELCTLKHAFEGSMKNLVKTIISGSPFVSLHYSDRLSVQLFK 240
Db 181 NNSSDIHALGCVLYELCTLKHAFEGSMKNLVKTIISGSPFVSLHYSDRLSVQLFK 240
QY 241 RNPRDRPSVNSILKGFIAKRIEKLSPOLIAEEFCLKTFESKFSQPIPAKRPASGNSI 300
Db 241 RNPRDRPSVNSILKGFIAKRIEKLSPOLIAEEFCLKTFESKFSQPIPAKRPASGNSI 300
QY 301 SVMPAQKITRPAKYGIPLAYKKYGDGKLHEKKPLQKHQAHOPEKRVNTGEEERKISE 360
Db 301 SVMPAQKITRPAKYGIPLAYKKYGDGKLHEKKPLQKHQAHOPEKRVNTGEEERKISE 360
```

```

OY 361 EAARRRLEFTEKROKODIISLMKAOMKROKRELERINAROGWRNVLASGSGE 420
DB 361 EAARRRLEFTEKROKODIISLMKAOMKROKRELERINAROGWRNVLASGSGE 420
OY 421 VKAPFLGSGGTAPSSFSRGOYEYHAIFDOMOQOARADNEAKWKREIYGRGLPEROKG 480
DB 421 VKAPFLGSGGTAPSSFSRGOYEYHAIFDOMOQOARADNEAKWKREIYGRGLPEROKG 480
OY 481 OLAVERAQOVEEFLORREKREAMONKARAGHMYLAPLROIRLONFNEROOIKAKLGEKK 540
DB 481 OLAVERAQOVEEFLORREKREAMONKARAGHMYLAPLROIRLONFNEROOIKAKLGEKK 540
OY 541 EANHSGOGSEPADMRKKTESLKAHANARAAYLKEOLEKREKREAVERRKWEHHLVA 600
DB 541 EANHSGOGSEPADMRKKTESLKAHANARAAYLKEOLEKREKREAVERRKWEHHLVA 600
OY 601 KGVKSSDVPSPLOHETGSGSPKQOMRSVITVTSALKRVGVDSSLTDTRETSEMOKTN 660
DB 601 KGVKSSDVPSPLOHETGSGSPKQOMRSVITVTSALKRVGVDSSLTDTRETSEMOKTN 660
OY 661 AISKREILRLNENKAKADEKMONLSDPFEINVEDAKENKESVSDRRKRWAGG 720
DB 661 AISKREILRLNENKAKADEKMONLSDPFEINVEDAKENKESVSDRRKRWAGG 720
OY 721 QLVIPDELTLDTSTETREHIVGEYIKLPGNSRRAMKGSPTDSYKILIGEAELQLOF 780
DB 721 QLVIPDELTLDTSTETREHIVGEYIKLPGNSRRAMKGSPTDSYKILIGEAELQLOF 780
OY 781 ELLENTTIRSEISPEGEKRYPLITGKRVKVOCSHFINPVAIVDSPEVETKSPSEASPOK 840
DB 781 ELLENTTIRSEISPEGEKRYPLITGKRVKVOCSHFINPVAIVDSPEVETKSPSEASPOK 840
OY 841 SLKLEGNLEPDDLETEILOEPESGINKKDESLCTITVDWISSEKTEKTOGADRITIOEN 900
DB 841 SLKLEGNLEPDDLETEILOEPESGINKKDESLCTITVDWISSEKTEKTOGADRITIOEN 900
OY 901 EVSEEGVSTVDOLSDIHIEPGTNDOSHKCDVKSVOPEPFHVKVSEHLNLPVOYS 960
DB 901 EVSEEGVSTVDOLSDIHIEPGTNDOSHKCDVKSVOPEPFHVKVSEHLNLPVOYS 960
OY 961 VQCSPEESFAPRSHSLPRKNKNKNSLLGLSTGLFDANPNKMLCTCSLPDLSKLFITLM 1020
DB 961 VQCSPEESFAPRSHSLPRKNKNKNSLLGLSTGLFDANPNKMLCTCSLPDLSKLFITLM 1020
OY 1021 DVPYVGVRODNLEIDETKDNITKEGPOSEDIYFERTTDOLOEIOASBOELRPOGEE 1080
DB 1021 DVPYVGVRODNLEIDETKDNITKEGPOSEDIYFERTTDOLOEIOASBOELRPOGEE 1080
OY 1081 YSEEEESVLKNSDYEPITANGTVADEEDNPSSSALNEBMSDNGELTASCECDSDVEN 1140
DB 1081 YSEEEESVLKNSDYEPITANGTVADEEDNPSSSALNEBMSDNGELTASCECDSDVEN 1140
OY 1141 HLEELRLHLEOEMGEFEFEYERIKAIHEDEENIEICSKIVONILGNEHQHLYAKLTH 1200
DB 1141 HLEELRLHLEOEMGEFEFEYERIKAIHEDEENIEICSKIVONILGNEHQHLYAKLTH 1200
OY 1201 LVMADGAYOEDNDE 1214
DB 1201 LVMADGAYOEDNDE 1214

```

```

; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 2355
; LENGTH: 1243
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-488-725A-2355

Query Match          99.5%; Score 6213; DB 18; Length 1242;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1213; Conservative 0; Mismatches 1; Indels 28; Gaps 1;

OY 1 MEKVRLQKIGEGSFGKAILVSTEDGROYVLIKEINISRMSSKREESREAVLANMKH 60
DB 1 MEKVRLQKIGEGSFGKAILVSTEDGROYVLIKEINISRMSSKREESREAVLANMKH 60
OY 61 PNIVQYRESEENGSLYIWMDCGEGDLFRRINAKGVLFQEDDILDMFOVIALKRVH 120
DB 61 PNIVQYRESEENGSLYIWMDCGEGDLFRRINAKGVLFQEDDILDMFOVIALKRVH 120
OY 121 DRKILHDIKSONFLTKGTVOLGDFGTAUVNSTVETLARTCGIPTYYSPEICEKRY 180
DB 121 DRKILHDIKSONFLTKGTVOLGDFGTAUVNSTVETLARTCGIPTYYSPEICEKRY 180
OY 181 NKSQDIALGCVLYETCLTKHAFEGSMKNVLKIIISGFPVSLHYSYDLRSIVSOLF 240
DB 181 NKSQDIALGCVLYETCLTKHAFEGSMKNVLKIIISGFPVSLHYSYDLRSIVSOLF 240
OY 241 RNPDRBPSVNSILEKGTIARIRKFLSPOLIAEECTTFSGFSOPIPAKRPASGNSI 300
DB 241 RNPDRBPSVNSILEKGTIARIRKFLSPOLIAEECTTFSGFSOPIPAKRPASGNSI 300
OY 301 SYMPAOKITKPAKYGIPLAYKRYGDKLHEKKPLQKHOAHQTEKRYNTGERRKISE 360
DB 301 SYMPAOKITKPAKYGIPLAYKRYGDKLHEKKPLQKHOAHQTEKRYNTGERRKISE 360
OY 361 EAARRRLEFTEKROKODIISLMKAOMKROKRELERINAROGWRNVLASGSGE 420
DB 361 EAARRRLEFTEKROKODIISLMKAOMKROKRELERINAROGWRNVLASGSGE 420
OY 421 VKAPFLGSGGTAPSSFSRGOYEYHAIFDOMOQOARADNEAKWKREIYGRGLPEROKG 480
DB 421 VKAPFLGSGGTAPSSFSRGOYEYHAIFDOMOQOARADNEAKWKREIYGRGLPEROKG 480
OY 481 OLAVERAQOVEEFLORREKREAMONKARAGHMYLAPLROIRLONFNEROOIKAKLGEKK 540
DB 481 OLAVERAQOVEEFLORREKREAMONKARAGHMYLAPLROIRLONFNEROOIKAKLGEKK 540
OY 541 YLAPLROIRLONFNEROOIKAKLGEKKANHSBGOSEADMRKKTESLKAHANARA 600
DB 541 YLAPLROIRLONFNEROOIKAKLGEKKANHSBGOSEADMRKKTESLKAHANARA 600
OY 601 AVLKEDLEKREKREAVERRKWEHHLVAKGVKSSDVPSPLOHETGSGSPKQOMRSVITV 660
DB 601 AVLKEDLEKREKREAVERRKWEHHLVAKGVKSSDVPSPLOHETGSGSPKQOMRSVITV 660
OY 633 TSAKREVGVDSSLTDTRETSEMOKTNNAISSKREILRLNENKAKADEKMONLSDPFE 692
DB 633 TSAKREVGVDSSLTDTRETSEMOKTNNAISSKREILRLNENKAKADEKMONLSDPFE 692

```

```

; RESULT 3
; US-09-488-725A-2355
; Sequence 2355, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseg Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21

```

```

Db 661 TSALKEVGVSSLDTRTSEEMOKTNNNAISSKREILRLNENIKAQOEDEKQNLSDTF 720
Qy 693 ETINVEDAKKEHEKESVSSDRKKWEAGOLVPLDELITLDTSPSTERTVGEVILGPN 752
Db 721 ETINVEDAKKEHEKESVSSDRKKWEAGOLVPLDELITLDTSPSTERTVGEVILGPN 780
Qy 753 GSPRRAMGKSPDVSVLKILGEAELOLOTELLENTIRSEISPEGEKYPKLITGEKKVQCI 812
Db 781 GSPRRAMGKSPDVSVLKILGEAELOLOTELLENTIRSEISPEGEKYPKLITGEKKVQCI 840
Qy 813 SHEINPSAIVDSPVETKSPESSEASPOMSLKLGNLEPPDLETELLOEPSTGNKDESP 872
Db 841 SHEINPSAIVDSPVETKSPESSEASPOMSLKLGNLEPPDLETELLOEPSTGNKDESP 900
Qy 873 CITTVWAISEKETKETOSADRTITQENEVSEGVSTVQOLSDIHIEGTNDSONSKCD 932
Db 901 CITTVWAISEKETKETOSADRTITQENEVSEGVSTVQOLSDIHIEGTNDSONSKCD 960
Qy 933 VDKSVQPEPFPHKVVHSEHLNLVPOVOSVOCSPSEESFAFRSHSHLPPKKNKNSLLIGLS 992
Db 961 VDKSVQPEPFPHKVVHSEHLNLVPOVOSVOCSPSEESFAFRSHSHLPPKKNKNSLLIGLS 1020
Qy 993 TGLFPANNPKMLRTCSLPLDSLKLFRITLMDVPTVGDVRODMLIETDEIKDENIKGPDSDSD 1052
Db 1021 TGLFPANNPKMLRTCSLPLDSLKLFRITLMDVPTVGDVRODMLIETDEIKDENIKGPDSDSD 1080
Qy 1053 IYFEETDIDLOELQASMEQLREOPGEYESSEESVILKNSDVEPTANGTVDADDDNPSS 1112
Db 1081 IYFEETDIDLOELQASMEQLREOPGEYESSEESVILKNSDVEPTANGTVDADDDNPSS 1140
Qy 1113 ESALNEEHSNDSOGEIASECECSVFNHLEELRLHEOMGFEKFEVEYERKATIHDE 1172
Db 1141 ESALNEEHSNDSOGEIASECECSVFNHLEELRLHEOMGFEKFEVEYERKATIHDE 1200
Qy 1173 DENIECKIYONILGNHOLYAKILHLVMAOGAYOEONDE 1214
Db 1201 DENIECKIYONILGNHOLYAKILHLVMAOGAYOEONDE 1242

```

```

RESULT 4
US-09-783-320-6
; Sequence 6, Application US/09783320
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hu, Yi
; APPLICANT: Nepomniichy, Boris
; APPLICANT: Turner, C. Alexander Jr
; APPLICANT: Zambrowicz, Brian
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0137-USA
; CURRENT FILING DATE: 2001-02-15
; CURRENT APPLICATION NUMBER: US/09/783,320
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/184,014
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1007
; TYPE: PRY
; ORGANISM: homo sapiens
US-09-783-320-6

```

Query Match 82.7%; Score 5162; DB 21; Length 1007;
 Best Local Similarity 100.0%; Pred. No. 7.4e-273;
 Matches 1007; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 208 MNVLKTIISGSPVPSIHYSDLSVLSOLFKNPDRSIVSIIKGIARIKRIKPLS 267
Db 1 MNVLKTIISGSPVPSIHYSDLSVLSOLFKNPDRSIVSIIKGIARIKRIKPLS 60

```

```

Qy 268 POLIAEFCLKTPSPKRGSPPIPAKRPAASONSISVPAOKITTKPAANYGIPLAYKXYGDK 327
Db 61 POLIAEFCLKTPSPKRGSPPIPAKRPAASONSISVPAOKITTKPAANYGIPLAYKXYGDK 120
Qy 328 KLEHKKPLQKNOAOHOTPEKRVNTEGERRKISEEAAARRRLEFTEKKOKOQIISLMA 387
Db 121 KLEHKKPLQKNOAOHOTPEKRVNTEGERRKISEEAAARRRLEFTEKKOKOQIISLMA 180
Qy 388 EOMKROEKLEELINRARGOGRNVLSAGGSEVAPPLGSGGTIAPSFSRQYEHY 447
Db 181 EOMKROEKLEELINRARGOGRNVLSAGGSEVAPPLGSGGTIAPSFSRQYEHY 240
Qy 448 AITDQOQOAAEENKAKREITYGSLPRQKGLAVRAKAOVEPLORREMONKARA 507
Db 241 AITDQOQOAAEENKAKREITYGSLPRQKGLAVRAKAOVEPLORREMONKARA 300
Qy 508 EGHVYLARLQIRLONENRQOIKAKLGEKKEANHSEGOGESEADMRKKTIESLKAH 567
Db 301 EGHVYLARLQIRLONENRQOIKAKLGEKKEANHSEGOGESEADMRKKTIESLKAH 360
Qy 568 ANRAAVALKEOLERKREKREAYEREKWEHEHLVAKGVKSSDVSPLQGHETGSPSKOQMR 627
Db 361 ANRAAVALKEOLERKREKREAYEREKWEHEHLVAKGVKSSDVSPLQGHETGSPSKOQMR 420
Qy 628 SVISVTSALKEVGVSSLDTRTSEEMOKTNNNAISSKREILRLNENIKAQOEDEKGMON 687
Db 421 SVISVTSALKEVGVSSLDTRTSEEMOKTNNNAISSKREILRLNENIKAQOEDEKGMON 480
Qy 688 LSPTFEINVEDAKKEHEKESVSSDRKKWEAGOLVPLDELITLDTSPSTERTVGEVI 747
Db 481 LSPTFEINVEDAKKEHEKESVSSDRKKWEAGOLVPLDELITLDTSPSTERTVGEVI 540
Qy 748 KLEPNSSPRAMGKSPDVSVLKILGEAELOLOTELLENTIRSEISPEGEKYPKLITGER 807
Db 541 KLEPNSSPRAMGKSPDVSVLKILGEAELOLOTELLENTIRSEISPEGEKYPKLITGER 600
Qy 808 KVOCISHEINPSAIVDSPVETKSPESSEASPOMSLKLGNLEPPDLETELLOEPSTGN 867
Db 601 KVOCISHEINPSAIVDSPVETKSPESSEASPOMSLKLGNLEPPDLETELLOEPSTGN 660
Qy 868 DESLPTITDVMISEKETKETOSADRTITQENEVSEGVSTVQOLSDIHIEGTNDSDQ 927
Db 661 DESLPTITDVMISEKETKETOSADRTITQENEVSEGVSTVQOLSDIHIEGTNDSDQ 720
Qy 928 HSKCDVDKSVQPEPFPHKVVHSEHLNLVPOVOSVOCSPSEESFAFRSHSHLPPKKNKNSL 987
Db 721 HSKCDVDKSVQPEPFPHKVVHSEHLNLVPOVOSVOCSPSEESFAFRSHSHLPPKKNKNSL 780
Qy 988 LIGLSTGLFPANNPKMLRTCSLPLDSLKLFRITLMDVPTVGDVRODMLIETDEIKDENIKGP 1047
Db 781 LIGLSTGLFPANNPKMLRTCSLPLDSLKLFRITLMDVPTVGDVRODMLIETDEIKDENIKGP 840
Qy 1048 SDSEDIYFEETDIDLOELQASMEQLREOPGEYESSEESVILKNSDVEPTANGTVDADDD 1107
Db 841 SDSEDIYFEETDIDLOELQASMEQLREOPGEYESSEESVILKNSDVEPTANGTVDADDD 900
Qy 1108 DNSSSALNEEHSNDSOGEIASECECSVFNHLEELRLHEOMGFEKFEVEYERKATIH 1167
Db 901 DNSSSALNEEHSNDSOGEIASECECSVFNHLEELRLHEOMGFEKFEVEYERKATIH 960
Qy 1168 IHEDEDENEICKIYONILGNHOLYAKILHLVMAOGAYOEONDE 1214
Db 961 IHEDEDENEICKIYONILGNHOLYAKILHLVMAOGAYOEONDE 1007

```

```

RESULT 5
US-09-783-320-2
; Sequence 2, Application US/09783320
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hu, Yi
; APPLICANT: Nepomniichy, Boris
; APPLICANT: Turner, C. Alexander Jr

```

```

; APPLICANT: Zambrowicz, Brian
; FILE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
; TITLE REFERENCE: LEX-0137-05A
; CURRENT APPLICATION NUMBER: US/09/783,320
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: US 60/183,582
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/184,014
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1035
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-783-320-2

```

```

Query Match      82.3% Score 5138; DB 21; Length 1035;
Best Local Similarity 97.3%; Pred. No. 1.6e-271;
Matches 1007; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

OY 208 MKNVLTITSGSPVSLHSHYDLSVLSQLEKRNPRPRPSVNSLLEKGTAKRIEKLPS 267
      |||||||
DB 1 MKNVLTITSGSPVSLHSHYDLSVLSQLEKRNPRPRPSVNSLLEKGTAKRIEKLPS 60

OY 268 POLIAEEFLKTSKSGSOPIPAKRPASGONSTVMPAKITKPAKYGIPLAYKKYGDK 327
      |||||||
DB 61 POLIAEEFLKTSKSGSOPIPAKRPASGONSTVMPAKITKPAKYGIPLAYKKYGDK 120

OY 338 KLHEKKPLQKHQAHTPEKRVNTGERRKISEAARKRRLFEIEKKOKDOIIISLKA 387
      |||||||
DB 121 KLHEKKPLQKHQAHTPEKRVNTGERRKISEAARKRRLFEIEKKOKDOIIISLKA 180

OY 388 EOMKROEKERLERINRAEOGMWNVLSAGSGEYKAPFLGSGTITAPSSFSRGYEYH 447
      |||||||
DB 181 EOMKROEKERLERINRAEOGMWNVLSAGSGEYKAPFLGSGTITAPSSFSRGYEYH 240

OY 448 AIFDMMOQRAEDNEAKWKREIYGRGLPEROKGQLAVEAKQVEEFLQKKREAMONKRA 507
      |||||||
DB 241 AIFDMMOQRAEDNEAKWKREIYGRGLPEROKGQLAVEAKQVEEFLQKKREAMONKRA 300

OY 508 EGHM-----YLAIRLQRIQNFNEQOIKAKIRGEX 539
      |||||||
DB 301 EGHMIIQNLAMYGGRSSRGKPRNKEEYVYLAIRLQRIQNFNEQOIKAKIRGEX 360

OY 540 KEANSEGOEGSEEDMRKKIESTLKHANARAAYLKEOLERRKREAYEREKVEEHLY 599
      |||||||
DB 361 KEANSEGOEGSEEDMRKKIESTLKHANARAAYLKEOLERRKREAYEREKVEEHLY 420

OY 600 AKGVSSVSPRLGOHETGSGSKOQMRVISTSAKKEGVDSLTDRRETSSEMOKTN 659
      |||||||
DB 421 AKGVSSVSPRLGOHETGSGSKOQMRVISTSAKKEGVDSLTDRRETSSEMOKTN 480

OY 660 MAISSKREILRLNENLKAODEGKMONLSDTFEIVNHEDAKHEKEKSVSDRKKWEG 719
      |||||||
DB 481 MAISSKREILRLNENLKAODEGKMONLSDTFEIVNHEDAKHEKEKSVSDRKKWEG 540

OY 720 GOLVLPDLDELDTISSTTERHNVGEYIKLGPNGSPRRAMGKPTVSYLKILGEALDQ 779
      |||||||
DB 541 GOLVLPDLDELDTISSTTERHNVGEYIKLGPNGSPRRAMGKPTVSYLKILGEALDQ 600

OY 780 TELLENTTIRSEISPEGEKRYPLITGEEKVOCISHETNPASIVDSVETKSPFESASQ 839
      |||||||
DB 601 TELLENTTIRSEISPEGEKRYPLITGEEKVOCISHETNPASIVDSVETKSPFESASQ 660

OY 840 MSILKRGNLSEEDDLETEILOPESGTNKESTPCTITDVAIISEKETKETOSADRTIOE 899
      |||||||
DB 661 MSILKRGNLSEEDDLETEILOPESGTNKESTPCTITDVAIISEKETKETOSADRTIOE 720

OY 900 NEVSEGVSSVTDOLSDIHIEGPTNDSQSKCDVAKVQEPFPHKVVHSEHLNLPVOY 959
      |||||||
DB 721 NEVSEGVSSVTDOLSDIHIEGPTNDSQSKCDVAKVQEPFPHKVVHSEHLNLPVOY 780

```

```

OY 960 SVQSPSESPAFRSHSHLPPKNNKNNLSLIGSLGTFDANNPKMLRTCSLPDISKLEFRL 1019
      |||||||
DB 781 SVQSPSESPAFRSHSHLPPKNNKNNLSLIGSLGTFDANNPKMLRTCSLPDISKLEFRL 840

OY 1020 MDVPTVGDYKODMLEIDEIKDENIKGPESDSDIYEETDLDLOELQASMEQLLREQPGE 1079
      |||||||
DB 841 MDVPTVGDYKODMLEIDEIKDENIKGPESDSDIYEETDLDLOELQASMEQLLREQPGE 900

OY 1080 EYSEESRYLKNDSVPTANGTDVADEDNPSSESALNEEWHNSDGEIASECECDVSF 1139
      |||||||
DB 901 EYSEESRYLKNDSVPTANGTDVADEDNPSSESALNEEWHNSDGEIASECECDVSF 960

OY 1140 NHLEELRLHEDEKGEKFEVEYEKTAIHEDENIEICSKYONTIGNHOLYAKIL 1199
      |||||||
DB 961 NHLEELRLHEDEKGEKFEVEYEKTAIHEDENIEICSKYONTIGNHOLYAKIL 1020

OY 1200 HLYMADGAYOEDNDE 1214
      |||||||
DB 1021 HLYMADGAYOEDNDE 1035

RESULT 6
US-09-488-725A-5927
; Sequence 5927, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784F.PCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PC_Fl_genes_b Versions 1.0
; SEQ ID NO 5927
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-488-725A-5927

```

```

Query Match      75.9% Score 4735.5; DB 18; Length 966;
Best Local Similarity 96.3%; Pred. No. 1.3e-249;
Matches 930; Conservative 2; Mismatches 5; Indels 29; Gaps 2;

OY 278 KTFKSGSOPIT-PAKRPSGONSTVMPAKITKPAKYGIPLAYKKYGDKILHEKKPLQ 336
      |||||||
DB 1 KTFKSGSOPIT-PAKRPSGONSTVMPAKITKPAKYGIPLAYKKYGDKILHEKKPLQ 60

OY 337 KKHQAHTPEKRVNTGERRKISEAARKRRLFEIEKKOKOIIISLKAODMKROKE 396
      |||||||
DB 61 KKHQAHTPEKRVNTGERRKISEAARKRRLFEIEKKOKOIIISLKAODMKROKE 120

OY 397 RLEIRINAREOGWNVLSAGSGEYKAPFLGSGTITAPSSFSRGYEYHNAIFDMMOQ 456
      |||||||
DB 121 RLEIRINAREOGWNVLSAGSGEYKAPFLGSGTITAPSSFSRGYEYHNAIFDMMOQ 180

OY 457 RAEDNEAKWKREIYGRGLPEROKGQLAVEAKQVEEFLQKKREAMONKRAEBSHM----- 511

```

```

Db 181 RAEDNEAKMREIYGRGLPEROKGOLAVERAKQVEEFLOKRREBAONKRAFAGHIGLON 240
QY 512 -----VYLARIQIRLQNFNRQOIKAKLNGEKEKKAHNSGQ 548
Db 241 LAAMYGRRSSRGKPRNKEEYVYLARIQIRLQNFNRQOIKAKLNGEKEKKAHNSGQ 300
QY 549 EGSEADMRKRKTESLKAHANARAVALKEQLERKREKREKWEHLVAKGVSSDV 608
Db 301 EGSEADMRKRKTESLKAHANARAVALKEQLERKREKREKWEHLVAKGVSSDV 360
QY 609 SPPLGQHEHTGSPSKOQMRSVISVTSALKVGVDSLTDTRETSEMOKTNNATISSKREI 668
Db 361 SPPLGQHEHTGSPSKOQMRSVISVTSALKVGVDSLTDTRETSEMOKTNNATISSKREI 420
QY 669 LRLNENLKAQDEKQKQNLSDTFEINVEDAKHEKESVSDRKKWAGQOLVPLDE 728
Db 421 LRLNENLKAQDEKQKQNLSDTFEINVEDAKHEKESVSDRKKWAGQOLVPLDE 480
QY 729 LTLDTSEFTEHRTVGEVIKLGPNGSPRRAMGKSPYDSVLKILGEAELOLOTELENTTI 788
Db 481 LTLDTSEFTEHRTVGEVIKLGPNGSPRRAMGKSPYDSVLKILGEAELOLOTELENTTI 540
QY 789 RSEISPEGEKYPPLITGEKKVOCISHEINPSAIVDSPVETKSPFESEASPOMSLKLEGNL 848
Db 541 RSEISPEGEKYPPLITGEKKVOCISHEINPSAIVDSPVETKSPFESEASPOMSLKLEGNL 600
QY 849 EEPDDETEILOEPSGTNKDESLPCTITDVIWISSEKETKETOSADRITTOENEVEDGVS 908
Db 601 EEPDDETEILOEPSGTNKDESLPCTITDVIWISSEKETKETOSADRITTOENEVEDGVS 660
QY 909 STVQOLSDIHEPTNDSOHSKCDVKSVOPEPFKVVHSEHLNVPOVOSVOCSPSES 968
Db 661 STVQOLSDIHEPTNDSOHSKCDVKSVOPEPFKVVHSEHLNVPOVOSVOCSPSES 720
QY 969 FAFRSHSHLPKKNKNSLLIGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMQVPTVGDV 1028
Db 721 FAFRSHSHLPKKNKNSLLIGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMQVPTVGDV 780
QY 1029 RQDNLDEIDETIKENIKGSPDSSEDIYFEETDLOELQASMEQLREOPGESESESSV 1088
Db 781 RQDNLDEIDETIKENIKGSPDSSEDIYFEETDLOELQASMEQLREOPGESESESSV 840
QY 1089 LKNSDVEPTANGTDVADEDNPSSESALNEMHSDNSDGEIASSECECDSVFNLLELRH 1148
Db 841 LKNSDVEPTANGTDVADEDNPSSESALNEMHSDNSDGEIASSECECDSVFNLLELRH 900
QY 1149 LEQEMGEKPFVEYERIKAIHEDEDENIETISKIVONITGNEHOHLXAKILHLVADGAY 1208
Db 901 LEQEMGEKPFVEYERIKAIHEDEDENIETISKIVONITGNEHOHLXAKILHLVADGAY 960
QY 1209 QEDNDE 1214
Db 961 QEDNDE 966

```

```

; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO: 5928
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-5928

Query Match 75.9%; Score 4735.5; DB 18; Length 966;
Best Local Similarity 96.3%; Pred. No. 1,36-249;
Matches 930; Conservative 2; Mismatches 5; Indels 29; Gaps 2;

QY 278 KTFSEKSGSOPi-PAKRPASQNSISVMPAOKITTKPAKYGIPLAYKKYGDKKLHEKKPLQ 336
Db 1 KHFSKSGSALYOLKRPASQNSISVMPAOKITTKPAKYGIPLAYKKYGDKKLHEKKPLQ 60
QY 337 KHKQAHQTEPEKRYNVTGEERKISEBAARRRLFEETEKKKQKQDILSLMAEQMKROEKE 396
Db 61 KHKQAHQTEPEKRYNVTGEERKISEBAARRRLFEETEKKKQKQDILSLMAEQMKROEKE 120
QY 397 RLERIRAREQGRNVLASGSGEYKAPFLGSGGTIAPSSFSRGQYEHYHAIPDQMOQ 456
Db 121 RLERIRAREQGRNVLASGSGEYKAPFLGSGGTIAPSSFSRGQYEHYHAIPDQMOQ 180
QY 457 RAEDNAKAKKRELYGGLPEROKGOLAVERAKQVEEFLORREMOKARAREGMM----- 511
Db 181 RAEDNAKAKKRELYGGLPEROKGOLAVERAKQVEEFLORREMOKARAREGMM----- 240
QY 512 -----VYLARIQIRLQNFNRQOIKAKLNGEKEKKAHNSGQ 548
Db 241 LAAMYGRRSSRGKPRNKEEYVYLARIQIRLQNFNRQOIKAKLNGEKEKKAHNSGQ 300
QY 549 EGSEADMRKRKTESLKAHANARAVALKEQLERKREKREKWEHLVAKGVSSDV 608
Db 301 EGSEADMRKRKTESLKAHANARAVALKEQLERKREKREKWEHLVAKGVSSDV 360
QY 609 SPPLGQHEHTGSPSKOQMRSVISVTSALKVGVDSLTDTRETSEMOKTNNATISSKREI 668
Db 361 SPPLGQHEHTGSPSKOQMRSVISVTSALKVGVDSLTDTRETSEMOKTNNATISSKREI 420
QY 669 LRLNENLKAQDEKQKQNLSDTFEINVEDAKHEKESVSDRKKWAGQOLVPLDE 728
Db 421 LRLNENLKAQDEKQKQNLSDTFEINVEDAKHEKESVSDRKKWAGQOLVPLDE 480
QY 729 LTLDTSEFTEHRTVGEVIKLGPNGSPRRAMGKSPYDSVLKILGEAELOLOTELENTTI 788
Db 481 LTLDTSEFTEHRTVGEVIKLGPNGSPRRAMGKSPYDSVLKILGEAELOLOTELENTTI 540
QY 789 RSEISPEGEKYPPLITGEKKVOCISHEINPSAIVDSPVETKSPFESEASPOMSLKLEGNL 848
Db 541 RSEISPEGEKYPPLITGEKKVOCISHEINPSAIVDSPVETKSPFESEASPOMSLKLEGNL 600
QY 849 EEPDDETEILOEPSGTNKDESLPCTITDVIWISSEKETKETOSADRITTOENEVEDGVS 908
Db 601 EEPDDETEILOEPSGTNKDESLPCTITDVIWISSEKETKETOSADRITTOENEVEDGVS 660
QY 909 STVQOLSDIHEPTNDSOHSKCDVKSVOPEPFKVVHSEHLNVPOVOSVOCSPSES 968
Db 661 STVQOLSDIHEPTNDSOHSKCDVKSVOPEPFKVVHSEHLNVPOVOSVOCSPSES 720
QY 969 FAFRSHSHLPKKNKNSLLIGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMQVPTVGDV 1028
Db 721 FAFRSHSHLPKKNKNSLLIGLSTGLFDANNPKMLRTCSLPDLSKLFRTLMQVPTVGDV 780

```

```

RESULT 7
US-09-488-725A-5928
; Sequence 5928, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784PLCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19

```



```

10
US-09-873-404-4
Sequence 4, Application US/09873404
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THERIOF
FILE REFERENCE: CL001212-CIP
CURRENT APPLICATION NUMBER: US/09/873,404
CURRENT FILING DATE: 2001-06-05
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0.
SEQ ID NO 4
LENGTH: 616
TYPE: PRRT
ORGANISM: Mouse
US-09-873-404-4
Db 601 KGVKSSVSPPLQGHETG 618
Db 540 EADDTKQGEETLEETDRLKMKESLAQTNARAAVYLKOLEKKREKAEYERKKWEEHLYA 599
Db 480 HLAVERANQVEEFLQKRRLEAMONKRAEGHWYTLARLQRIPLQNFNPROOIKKLGCEK 539
Db 481 OLAVERRAKOYEETLQKRRLEAMONKRAEGHWYTLARLQRIPLQNFNPROOIKKLGCEK 540
Db 420 VKASFEGIGGAVSFPCSPRGQYEYHAIIPDMQRLRAEDNEAMKCGIYGRLLPERQK 479
Db 421 VKAFELGSGTIPSPFSRSGQYEYHAIIPDMOQRAEDNEAMKCKEYIGRLPERQK 480
Db 422 VKAFELGSGTIPSPFSRSGQYEYHAIIPDMOQRAEDNEAMKCKEYIGRLPERQK 480

```

Query Match	Similarity	44.6%	Score 2785	DB 22	Length 616
Best Local	Similarity	88.0%	Pred. No. 2,2e-143		
Matches	544	Conservative	27	Mismatches	45
				Indels	2
				Gaps	2
QY	1	MEKYVRLQKIGSESSFKALIVKSTEDGQYAIKEINISRMSSKREESREVVAVLNMKH	60		
Db	1	MEKYVRLQKIGSESSFKALIVKSTEDGHHYAIKEINISRMSSDKRQSSREVVAVLNMKH	60		
QY	61	PNIVQYRESPEENGSLYIVMDCEGGDLFPRKINMOKGVLEOEOIILMFQIOLAKIHH	120		
Db	61	PNIVQYKESPEENGSLYIVMDCEGGDLFPRKINMOKGALFQEOOILDMFQIOLAKIHH	120		
QY	121	DRKILHRDIKSONIPLTKDGYQLDQDFIAVLNISTVELARCTGTPYIYSLPEICENKRY	180		
Db	121	DRKILHRDIKSONIPLTKDGYQLDQDFIAVLNISTVELARCTGTPYIYSLPEICENKRY	180		
QY	181	NNNSDIALGAGVLYELCTLKHAFEGSMKNVLTIGSPSPSYAHSYQLRELYSQLR	240		
Db	181	NNNSDIALGAGVLYELCTLKHAFEGSMKNVLTIGSPSPSYAHSYQLRELYSQLR	240		
QY	241	RNRDPRSVNLSLEKFIKRIEKLSPQLIAEERCLAKTSKQSPPIPKKPAQSONSI	300		
Db	241	RNRDPRSVNLSLEKFIKRIEKLSPQLIAEERCLAKTSKQSPPIPKKPAQSQGVS	300		
QY	301	SVNPAQKIIKPAKAYIPLATKYKIDGDKLHEKKPLQKHQAQHPTEKRVNTGEERRKISE	360		
Db	301	SEVPAQKIIKPAKAYIPLATKYKIDGDKLHEKKPPKKAQHQIPVKKMNSGERKKRMBE	360		
QY	361	EAARKRRLEFIEKKKKOKDQIISLKAADMRKQERLERLINRAEDGMNVLSAGSGCE	420		
Db	361	EAARKRRLEFIEKKKKOKDO -IRFLKAQMRKQERLERINRRREGGMNVLRAGSGGE	419		
QY	421	VKAPLGSGGTIAPSFSSRGQYEHYHAIFDOMOOARADENAKMKREIYGRGLPRROG	480		
Db	420	VKASFQIGGAVSPSCSPRGQYEHYHAIFDOMOURLRAEDENAKMKQIGTGMPLPRROG	479		

QY	481	QVVEBAKQVVEEFLORKEEMANQKARAEHVYLAFLROIRLÖNNNEQOIKALREKK	54.0
Db	480	HLAVEANQVVEEFLORKEEMANQKARAEHVYLAFLROIRLÖNNNEQOIKALREKK	53.9
QY	541	EAAHSGQSGSEALMRKKLESLKAAHNAARVALEOLERKKAREKKVMEHLVA	600
Db	540	EALGTGQGTATEETMRKKMESLKAOTNAAARVALEOLERKKAREKKVMEHLVA	599
QY	601	KGVKSDVSPFLQGHETG	618
Db	600	R-VKSSDVPLPLELLETG	616

```

RESULT 11
US-09-870-962-6
Sequence 6, Application US/09870962
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Gorgone, Gina
APPLICANT: Azimzai, Valda
APPLICANT: Lu, Alina
TITLE OF INVENTION: Protein Kinase Homologs
FILE REFERENCE: PF-0614 US
CURRENT APPLICATION NUMBER: US/09/870,962
CURRENT FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/420,915
PRIOR FILING DATE: 1999-10-20
PRIOR APPLICATION NUMBER: US 09/173,581
PRIOR FILING DATE: 1998-10-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PERL Program
SEQ ID NO 6
LENGTH: 345
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: -
OTHER INFORMATION: 1567782
US-09-870-962-6

```

[illegible]

Wed May 15 14:13:57 2002

us-09-783-320-4, rapm

Page 10

Sequence 1590, Application US/09760446A
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PTZ68
CURRENT APPLICATION NUMBER: US/09/760,446A
CURRENT FILING DATE: 2000-01-16
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207

Wed May 15 14:13:57 2002

us-09-783-320-4.ram

Page 12

Db 504 ----LQNTKESKNPEOK---YKAKKGVKFEINLDKCIDENILOEEBAMIPNETLTFE 555
QY 694 ----INVHEDAKHEHEKESVSDDRKKW-EAGGOLVPIPLDELTLDTSF-----STERH 741
Db 556 DGMKFEYECVKEHGDYTDKAFKELHCPAAFTLTLWLSFLFLFKSLPHFLLEKSPFSRH 615
QY 742 TVGEVIRKLGPNKSPRAMKSPDVSUKIIGEAELOLOT 780
Db 616 LIEDLL-----CANDCSLKDMSKEMELRT 640

Search completed: May 15, 2002, 07:59:35
Job time: 395 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 06:30:42 ; Search time 128.31 Seconds
(without alignments)

1050.920 Million cell updates/sec

Title: US-09-783-320-4

Perfect score: 6243

Sequence: 1 MEKTVRLQKIGESFGKAIL.....YAKILHLVMDAGYQEDNDE 1214

Scoring table: GAPDP 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_033802.*

- 1: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT.*
- 2: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT.*
- 3: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT.*
- 4: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT.*
- 5: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT.*
- 6: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT.*
- 7: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT.*
- 8: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT.*
- 9: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT.*
- 10: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT.*
- 11: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT.*
- 12: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT.*
- 13: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT.*
- 14: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT.*
- 15: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT.*
- 16: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT.*
- 17: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT.*
- 18: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT.*
- 19: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT.*
- 20: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT.*
- 21: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT.*
- 22: /net/abs06/SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6243	100.0	1214	AAU07102	Human novel human
2	6237	99.9	1214	AAU07102	Human polypeptide
3	6213	99.5	1242	AAU07102	Human polypeptide
4	5162	82.7	1007	AAU07103	Human novel human
5	5138	82.3	1035	AAU07101	Human novel human
6	4735.5	75.9	966	AAU0996	Human polypeptide
7	4735.5	75.9	966	AAU0997	Human polypeptide
8	1734	27.8	345	AAU06753	Human protein kina
9	1734	27.8	345	AAU06753	Human protein kina
10	1387	22.2	320	AAU05221	Human protein sequ
11	1025	16.4	649	AAU03545	Human protein kina

12	882	14.1	506	22	AAU78344	Human protein SEQ
13	875	14.0	527	22	AAU79328	Human protein SEQ
14	806.5	12.9	510	21	AAU68778	Amino acid sequenc
15	628	10.1	658	22	AAU65702	Novel protein kina
16	604	9.7	645	22	AAU03546	Human protein kina
17	599.5	9.6	979	19	AAU40309	Human ITRK protein
18	599.5	9.6	979	21	AAU95294	Human GEF containi
19	593.5	9.4	836	22	AAU65703	Novel protein kina
20	584	9.5	841	22	AAU65338	Drosophila melanog
21	557	8.9	330	22	AAU92330	Human SER4 protein
22	550	8.8	330	22	AAU92330	Human NIK1 protein
23	538	8.6	336	21	AAU59148	Human NERK protein
24	519	8.3	360	21	AAU59147	Drosophila melanog
25	503	8.1	360	21	AAU59147	E. nidulans NIMA p
26	502.5	8.0	1142	22	AAU70713	S. cerevisiae apopt
27	494.5	7.9	1298	22	AAU68222	Amino acid sequenc
28	494.5	7.9	1353	22	AAU68219	Amino acid sequenc
29	493.5	7.9	1202	21	AAU08521	A murine Ste20-rel
30	492	7.9	1339	22	AAU78421	Human protein SEQ
31	492	7.9	1385	22	AAU79405	Human protein SEQ
32	490.5	7.9	1306	22	AAU68218	Amino acid sequenc
33	490	7.8	1269	22	AAU68223	Human protein kina
34	490	7.8	1360	21	AAU85263	Large NIK-Related
35	487.5	7.8	1360	22	AAU50059	Amino acid sequenc
36	487.5	7.8	1324	22	AAU68221	Human cardiovascu
37	486	7.8	1277	22	AAU68220	Human TGF-beta rec
38	485.5	7.8	302	21	AAU59143	Human TGF-beta rec
39	485.5	7.8	302	21	AAU59143	Human TGF-beta rec
40	485.5	7.8	302	21	AAU71956	Human TGF-beta rec
41	485.5	7.8	311	22	AAU75475	Human TGF-beta rec
42	485.5	7.7	1332	22	AAU68217	Amino acid sequenc
43	478.5	7.7	268	21	AAU59146	Human TGF-beta rec
44	478.5	7.7	302	21	AAU59146	Murine Dendritic C
45	477.5	7.6	313	21	AAU22799	Human nek4 protein

ALIGNMENTS

RESULT	1
AAU07102	standard: Protein: 1214 AA.
ID	AAU07102
AC	AAU07102
XX	24-OCT-2001 (first entry)
DT	Human novel human protein, NHP #2.
DE	Human novel human protein, NHP; breast cancer; prostate cancer;
XX	Human; novel human protein; NHP; breast cancer; prostate cancer;
KW	Immunogen; antibody; gene therapy; antisense.
XX	
OS	Homo sapiens.
XX	
PN	WO200161016-A2.
PD	23-AUG-2001.
XX	
PF	15-FEB-2001; 2001WO-US05356.
XX	
PR	18-FEB-2000; 2000US-0183582.
PR	22-FEB-2000; 2000US-0184014.
XX	
PA	(LEXI-) LEXICON GENETICS INC.
XX	
FI	Walke DW, Hu Y, Nepomichy B, Turner CA, Zambrowicz B;
XX	
XX	WPI; 2001-602793/55.
DR	N-PSDB; AAS11558.
XX	
PT	Isolated nucleic acids encoding novel human proteins useful for the
PT	treatment of disease and as probes for testing and detection -
XX	

PS Claim 4: Page 37-39; 69pp; English.

XX The invention relates to novel human proteins (NHP) and the nucleic acids encoding them. The nucleic acids encode mammalian transporter proteins and are useful for the treatment (e.g. by gene therapy or antisense technology) of any of a wide variety of symptoms associated with biological disorders (e.g. breast and prostate cancer) or imbalances as probes for the identification, selection and validation of novel molecular targets for drug discovery. The proteins may be used to raise anti-NHP antibodies. The present sequence represents an NHP of the invention.

XX Sequence 1214 AA:

Query Match 100.0%; Score 6243; DB 22; Length 1214;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYVRLQKIGSGSPGKAILVKSTEDGRQVYIKENISRMSSKERESEREVAVLANMKH 60
DB 1 mekyvrlqkigsgsfgkailvkstedrqrvyikeinismsskereesrevavlanmhk 60
QY 61 PNIIVYRESFEENGSLYIVMDYCEGGDLFRINAKGVLFQEDQILDMEVQICLAKHYH 120
DB 61 pniivyresfeengslyivmdyceggdlfrinaqgvlfqedqildmfvqiclaikhvh 120
QY 121 DKKIIRHDTKSONIFLTGDTVOLGPGIARVLNSTVLELARTGICMPYILSPETCENKPY 180
DB 121 dkkilhrdtksonifltgdgtvolgpgiarvlnstvelartcigmpyylspelcenkpy 180
QY 181 NNRSDIMALGVLYELCTLKAFAFGSMKNLYIKISGSPPSIHYSLBLSVQLTK 240
DB 181 nnrsdimalgvlyelctlkafafgsmknlyikisgsppsihyslbrslvqltk 240
QY 241 RNPDRPSVNSTLEKGFIAKRIEKLSPOLIAEFCLKTSGKSGOPTPAKRPAQSONSI 300
DB 241 rnpdrpsvnsstlekgfiakrieklspoliaefclktsgksgoptpakrapaqsonsi 300
QY 301 SWPAQKITTPAKAAGIPLAVKRYGDKRLHEKKPLQKHKAQHOATPEKRVMTGEERRISE 360
DB 301 swpaqkiltkpaakgiplavkrygdkrlhekplqkhkqhahqpekrvtlgeerrikise 360
QY 361 EAARRRRLFEIEKEKKQDQIISLMKAQOMRKQEKRLERINRAEQGMHNVLSAGSGS 420
DB 361 eaarrrrlefeiekekkgdqilislmkaeqmkrqeklerinraeqgmvnvlisagsgs 420
QY 421 VKAPFGSGGTAPSSFSRQGYEHYHIFDOMQOARADNEAKWKKEIYGRGIPEROKG 480
DB 421 vkapfsgsgtapsfsfsgryehyhifdomqoaradneakwkkeiygrgiperpqkg 480
QY 481 OLAVERAKOVEEFLORKEAMONKARAGHMYLARTROLQNFNROOIRAKLGEKK 540
DB 481 olaverakoveeflqrkeamgnkareeghmvyarlrlqnlfnrgeqikakllygekk 540
QY 541 EANHSEGOGESEADMRKKKTESLKAHANAAVALKQLEKREKREAVERRKAKWEEHYLA 600
DB 541 eanhsegggeeeadmrrkkkleslkahanaavalkleqrkearekkrkweehylla 600
QY 601 KGVKSDVSPPLQGHETGSGPSKQOMRSVISTSALEKGVGDSLTDTRETSEMOXTNN 660
DB 601 kgvksdvspplqghetgsgpskqomrsvisvtalkevgydsalltltretseemqktnn 660
QY 661 AATSKREILRLNENIKAQEDEKGMQLSDTFEINVEDHKKHEHEKESVSSDRKKKWEAG 720
DB 661 aatsskreilrlnenikaqedeqgmqlsdtfelnvedahhehekesvsddckkweag 720
QY 721 QVIVPLDELTLDTSTSTERTVGVKILGPNGSPRRAMGKSPDVSYLKLTIGAELQLOLQ 780
DB 721 qvivpldeltltdststertlvgvkiligpnspramgkspdvslklkltigaellqlq 780
QY 781 ELLENTTTRSETSPGEGYKPLINGEKKVOCISHEINPAILVOSPVTETKSPESSESPOM 840
DB 781 ellenttrsetspgegykplingekkvocisheinpailvosptetkspesseqspom 840

DB 781 ellenttrsetspgegykplingekkvocisheinpailvosptetkspesseqspom 840
QY 841 SLKLEGNLEPPDDLEPEFIILOEPSTGNKDESLPCTTDTDWISSEKRTKETSADRTIOEN 900
DB 841 slklegnleppddleleilqepstgnkdeslpcttdtdwissekretketqadrtiigen 900
QY 901 EYSDGVSSVVDLSIDHITPEGNDOSHKCPVDKSVQPEPFHKVYVSEHLNLYPOVQS 960
DB 901 eysdgvssvvdlsidhitpegnoskcpvdksvqpepfhkvvsehlnlypvqqs 960
QY 961 VQCSPEESFAPRSHSLPFRKNKNLIGLSTGCFDANPNMFLTCSLPDLSTKFLTM 1020
DB 961 vqcspeesfaprsishslpfrknknlliglstgcfdanpnmfltlcslpdlstfltm 1020
QY 1021 DVPTGVDVRODNLLEIDEIKENIKESPSDEDTVEETDTDLQELQASMEQLLRPOGEE 1080
DB 1021 dvptgvdvrdnlleideikenikespsdesdtvteetdtldelqasmeqlireqpee 1080
QY 1081 YSEEESEVLKNSDVEPTANGTDVADDDNPPSSSALNEBMHSDNSGELIASECDESVFN 1140
DB 1081 yseeeesvlnksdveptangtdvadeddnpssesalnneebhdsndgeiaasecdsvfn 1140
QY 1141 HLEELRLHLEBOEMGFEPFVYEKIKAIHEDEDENEICSKIVONILGNEHQHLYAKILH 1200
DB 1141 hleelrlhleboemgfepfvyeikaihedeedenieicskivonilgnehqhlyakilh 1200
QY 1201 LYMADGAYORNDNE 1214
DB 1201 lymadgaygednde 1214

RESULT 2

AA039211 standard; Protein; 1214 AA.

AA039211:

22-OCT-2001 (first entry)

Human polypeptide SEQ ID NO 2356.

Human; noctropic; immunosuppressant; cytosolic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; leukemia; thrombolytic; drug screening; arthritis; inflammation;

Homo sapiens.

MO20015312-A1.

26-JUL-2001.

26-DEC-2000; 2000NC-US34263.

21-JAN-2000; 2000US-0488725.

25-APR-2000; 2000US-0532317.

09-JUL-2000; 2000US-0598042.

19-JUL-2000; 2000US-0620312.

03-AUG-2000; 2000US-0653450.

14-SEP-2000; 2000US-0662191.

19-OCT-2000; 2000US-0693036.

29-NOV-2000; 2000US-0727344.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J; Zhao QA, Zhou P, Goodrich R, Drmanac RT;

WPI: 2001-442253/47.

N-PSDB: AA158367.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

Example 4; SEQ ID NO 2356; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161365) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: Immune system suppression,
 CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

CC Sequence 1214 AA:

Query Match 99.9%; Score 6237; DB 22; Length 1214;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1213; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MEXYRLQKIGESFGKALIVKSTEDROVYKEINISRMSSRESREPRAYAVLANMKH 60
DB 1 MEXYRLQKIGESFGKALIVKSTEDROVYKEINISRMSSRESREPRAYAVLANMKH 60
QY 61 PRTIVYRSFEENSGSLIYMDYCGGDLFRINAKQGVLFQEQIILDMFVQICLAKYH 120
DB 61 PRTIVYRSFEENSGSLIYMDYCGGDLFRINAKQGVLFQEQIILDMFVQICLAKYH 120
QY 121 DKKILHDKSQNIFLTQDGVQDFGIARVLNLSVEYLARTCTGPPYVLSPEICENPY 180
DB 121 DKKILHDKSQNIFLTQDGVQDFGIARVLNLSVEYLARTCTGPPYVLSPEICENPY 180
QY 181 NNSKDIWALGCVLYELCTLKHAFAGSNKMLVLIISGSPPYSLHYSDLSRLVSQLEK 240
DB 181 NNSKDIWALGCVLYELCTLKHAFAGSNKMLVLIISGSPPYSLHYSDLSRLVSQLEK 240
QY 241 RMPRDRPSVNSILKEGFIAKRIEKFELSPQILAEFECLTKPSKGSOPIPAKPRASGNSI 300
DB 241 RMPRDRPSVNSILKEGFIAKRIEKFELSPQILAEFECLTKPSKGSOPIPAKPRASGNSI 300
QY 301 SYMPAKOKITTPAKKYGIPLAYKKYGGDKLHEKKRPLQKHQAHOPEKPVNTGSEERKISE 360
DB 301 SYMPAKOKITTPAKKYGIPLAYKKYGGDKLHEKKRPLQKHQAHOPEKPVNTGSEERKISE 360
QY 361 EAARPRULETTEKKEKKDOILIMKADOMKROKRELERINARROGMNRYLSAGSGE 420
DB 361 EAARPRULETTEKKEKKDOILIMKADOMKROKRELERINARROGMNRYLSAGSGE 420
QY 421 VAAPTLGSQGTIAESSSESSRCQYEHYHAIFPOMQOAAENNEKMKREITGRILPRKOG 480
DB 421 VAAPTLGSQGTIAESSSESSRCQYEHYHAIFPOMQOAAENNEKMKREITGRILPRKOG 480
QY 481 QLAVERAKOVEEFLQKREKADAMONKARAGHMYTLARLQIRLONFNENROOIRAKKLGEEK 540
DB 481 QLAVERAKOVEEFLQKREKADAMONKARAGHMYTLARLQIRLONFNENROOIRAKKLGEEK 540
QY 541 EANHSEGOEGSEADMRKKITESTLKAHANARAVALKQLEKRRKEAVERKKWBEHLVA 600
DB 541 EANHSEGOEGSEADMRKKITESTLKAHANARAVALKQLEKRRKEAVERKKWBEHLVA 600
QY 601 KGVKSDVSPPLGQHEHETGSPSKDOOMRSVITSYLSALKEVGVDSLTDTRETSEEMKOTNN 660
DB 601 KGVKSDVSPPLGQHEHETGSPSKDOOMRSVITSYLSALKEVGVDSLTDTRETSEEMKOTNN 660

```

```

QY 661 ATSSKREILRRNLNKAODEKMONLSDTFEINYHEDAKHEHEKRSVSDRRKMEAGG 720
DB 661 ATSSKREILRRNLNKAODEKMONLSDTFEINYHEDAKHEHEKRSVSDRRKMEAGG 720
QY 721 QLVPLDELDTLDTSPSTTERHTVEVTKLGPNKSPRANGKSPPTSVLKILGEALQLOT 780
DB 721 QLVPLDELDTLDTSPSTTERHTVEVTKLGPNKSPRANGKSPPTSVLKILGEALQLOT 780
QY 781 ELLENTTIRSEISPEGGEKRYLITGEEKKQOCISHEINPSAIVDSVYEMKSPFSRSPOM 840
DB 781 ELLENTTIRSEISPEGGEKRYLITGEEKKQOCISHEINPSAIVDSVYEMKSPFSRSPOM 840
QY 841 SKLEGNLLEPPDLTEITLQEPSTGNDESILPCTITDVMISEKKEKTEQOSADRTIGEN 900
DB 841 SKLEGNLLEPPDLTEITLQEPSTGNDESILPCTITDVMISEKKEKTEQOSADRTIGEN 900
QY 901 EYSEDEGVSTYDQSLDIHEFGTNDQSHSKCDVKSQVPEPFHKKVYSEHLNLVPOVS 960
DB 901 EYSEDEGVSTYDQSLDIHEFGTNDQSHSKCDVKSQVPEPFHKKVYSEHLNLVPOVS 960
QY 961 VOCSPSESPAFKSHSLPPKNNKNSLLIGLSTGLFDANNPKMLRTCSLPDLISKLFTLM 1020
DB 961 VOCSPSESPAFKSHSLPPKNNKNSLLIGLSTGLFDANNPKMLRTCSLPDLISKLFTLM 1020
QY 1021 DVPYGVDRQDNLFEIDELKIDENIKESGSDSDIYFEETDTLOEQAEMBOILRQPGEE 1080
DB 1021 DVPYGVDRQDNLFEIDELKIDENIKESGSDSDIYFEETDTLOEQAEMBOILRQPGEE 1080
QY 1081 YSESEESVYKNSDVEPTANGTDAVEDDNPSESALNEEHSNDSNGEIASRCECDSVFN 1140
DB 1081 YSESEESVYKNSDVEPTANGTDAVEDDNPSESALNEEHSNDSNGEIASRCECDSVFN 1140
QY 1141 HIEEIRLHEOMGFEKFEYEEKIKAIHEDEDENIEICSTYQNTLGNHGLYAKIIL 1200
DB 1141 HIEEIRLHEOMGFEKFEYEEKIKAIHEDEDENIEICSTYQNTLGNHGLYAKIIL 1200
QY 1201 LWADQAYQEDNDE 1214
DB 1201 LWADQAYQEDNDE 1214

```

RESULT 3
 AAM39210
 ID AAM39210 standard; Protein; 1242 AA.
 AC AAM39210;
 DT 22-OCT-2001 (first entry)
 XX Human polypeptide SEQ ID NO 2355.
 DE Human; nootropic; immunosuppressant; cytoskeletal; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW Leukemia.
 XX Homo sapiens.
 OS Homo sapiens.
 PN W0200153312-A1.
 PD 26-JUL-2001.
 XX 26-DEC-2000; 2000MO-US34263.
 XX 21-JAN-2000; 2000US-0488725.
 XX 25-APR-2000; 2000US-0535317.
 PR 09-JUL-2000; 2000US-0596042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR N-PSDB: AA158366.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 4: SEQ ID NO 2355: 10078bp: English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA158042-AA164213) with neurotropic,
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: Immune system suppression,
 CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX Sequence 1242 AA:
 SQ

Query Match 99.5%; Score 6213; DB 22; Length 1242;
 Best Local Similarity 97.7%; Pred. No. 0;
 Matches 1213; Conservative 0; Mismatches 1; Indels 28; Gaps 1;
 QY 1 MEKVRLOKTEGSGFKALIVKSTEDGRQVYKINISMSKSEKRESEKREAVANLAKKH 60
 DB 1 MEKVRLOKTEGSGFKALIVKSTEDGRQVYKINISMSKSEKRESEKREAVANLAKKH 60
 QY 61 PNIVQRESPEENGSLIVNDYCEGDLKFRINAKGVLEPDDQILDMFVQICLAKHYH 120
 DB 61 PNIVQRESPEENGSLIVNDYCEGDLKFRINAKGVLEPDDQILDMFVQICLAKHYH 120
 QY 121 DRKTLHNDIKSONIFLTKDGTVOLGDFGIARVLNSTVELARTICIGPYLLSPGICENKPY 180
 DB 121 DRKTLHNDIKSONIFLTKDGTVOLGDFGIARVLNSTVELARTICIGPYLLSPGICENKPY 180
 QY 181 NNSKDIWALACVLYELCTLKHAFEAGSMKNLVKITSFSPVSLVSDLSLVSQLFK 240
 DB 181 NNSKDIWALACVLYELCTLKHAFEAGSMKNLVKITSFSPVSLVSDLSLVSQLFK 240
 QY 241 RNPDRPSVNSIIEKGFIARIEKPLSPOLIAEFCLTFKSGKSGOPFARAKPAGSONSI 300
 DB 241 RNPDRPSVNSIIEKGFIARIEKPLSPOLIAEFCLTFKSGKSGOPFARAKPAGSONSI 300
 QY 301 SVMPAKOITKPAKAYGIPLAVKYGDKLHEKPKLOKHQAOTPEKRVNTEGEEPRKISE 360
 DB 301 SVMPAKOITKPAKAYGIPLAVKYGDKLHEKPKLOKHQAOTPEKRVNTEGEEPRKISE 360
 QY 361 EAARRRLEFTEKKKKKODIISLAKAOKKROEKRELERINAROGKRNVLASAGSGE 420
 DB 361 EAARRRLEFTEKKKKKODIISLAKAOKKROEKRELERINAROGKRNVLASAGSGE 420
 QY 421 VKAPFLSGGTIAPSFSFGQYEHYHAIFDQWQOQRAEDNEAKMKRELYTGGLEROKG 480
 DB 421 VKAPFLSGGTIAPSFSFSFGQYEHYHAIFDQWQOQRAEDNEAKMKRELYTGGLEROKG 480

QY 481 QLAVERAKQVEEFLOKRRKREAMONKARABGH-----V 512
 DB 481 QLAVERAKQVEEFLOKRRKREAMONKARABGH-----V 512
 QY 513 YLARLQIRLQNPENKQOIKAKLRGEKKPANSSEOGSEADMRKKRTESIAKANANARA 572
 DB 513 YLARLQIRLQNPENKQOIKAKLRGEKKPANSSEOGSEADMRKKRTESIAKANANARA 572
 QY 541 YLARLQIRLQNPENKQOIKAKLRGEKKPANSSEOGSEADMRKKRTESIAKANANARA 600
 DB 541 YLARLQIRLQNPENKQOIKAKLRGEKKPANSSEOGSEADMRKKRTESIAKANANARA 600
 QY 573 AVLKEQERKRPAYREKKVMEHLVAKGVSSDPSPPLGONHFGSGSPKQOORSVY 632
 DB 573 AVLKEQERKRPAYREKKVMEHLVAKGVSSDPSPPLGONHFGSGSPKQOORSVY 632
 QY 601 AVLKEQERKRPAYREKKVMEHLVAKGVSSDPSPPLGONHFGSGSPKQOORSVY 660
 DB 601 AVLKEQERKRPAYREKKVMEHLVAKGVSSDPSPPLGONHFGSGSPKQOORSVY 660
 QY 633 TSAKKEGVSSILTDRESEEMOKNNMISSEKREILRNENKLAODEKGMONLSPTF 692
 DB 633 TSAKKEGVSSILTDRESEEMOKNNMISSEKREILRNENKLAODEKGMONLSPTF 692
 QY 661 TSAKKEGVSSILTDRESEEMOKNNMISSEKREILRNENKLAODEKGMONLSPTF 720
 DB 661 TSAKKEGVSSILTDRESEEMOKNNMISSEKREILRNENKLAODEKGMONLSPTF 720
 QY 693 EINVEDAKHEKESVSDRKKMEAGOLVLPDEILDLTSFSTTEHNTYGEVYIKGPN 752
 DB 693 EINVEDAKHEKESVSDRKKMEAGOLVLPDEILDLTSFSTTEHNTYGEVYIKGPN 752
 QY 721 ELNVHEDAKHEKESVSDRKKMEAGOLVLPDEILDLTSFSTTEHNTYGEVYIKGPN 780
 DB 721 ELNVHEDAKHEKESVSDRKKMEAGOLVLPDEILDLTSFSTTEHNTYGEVYIKGPN 780
 QY 753 GSPRRWAKSPVSVLKLGEAELOLQTELLENTTIRSEISPEGEKYPRLITGEKKVQCI 812
 DB 753 GSPRRWAKSPVSVLKLGEAELOLQTELLENTTIRSEISPEGEKYPRLITGEKKVQCI 812
 QY 781 GSPRRWAKSPVSVLKLGEAELOLQTELLENTTIRSEISPEGEKYPRLITGEKKVQCI 840
 DB 781 GSPRRWAKSPVSVLKLGEAELOLQTELLENTTIRSEISPEGEKYPRLITGEKKVQCI 840
 QY 813 SHEINPSAIVDSPEVETKSPFSEASPOMSLKLEGNLEPDDLETELLOEPSTNKKDSLP 872
 DB 813 SHEINPSAIVDSPEVETKSPFSEASPOMSLKLEGNLEPDDLETELLOEPSTNKKDSLP 872
 QY 841 SHEINPSAIVDSPEVETKSPFSEASPOMSLKLEGNLEPDDLETELLOEPSTNKKDSLP 900
 DB 841 SHEINPSAIVDSPEVETKSPFSEASPOMSLKLEGNLEPDDLETELLOEPSTNKKDSLP 900
 QY 873 CTITDVAISEKTEKRTGOSADRTITIOENESVDGVSSTVDQSLDIHIEPTGNDQSOKCD 932
 DB 873 CTITDVAISEKTEKRTGOSADRTITIOENESVDGVSSTVDQSLDIHIEPTGNDQSOKCD 932
 QY 901 CLITDVAISEKTEKRTGOSADRTITIOENESVDGVSSTVDQSLDIHIEPTGNDQSOKCD 960
 DB 901 CLITDVAISEKTEKRTGOSADRTITIOENESVDGVSSTVDQSLDIHIEPTGNDQSOKCD 960
 QY 933 VDKSVQEPFPHKVVSEHNLTPVOVQVQSPESFARSHHPKNNKNSLILGLS 992
 DB 933 VDKSVQEPFPHKVVSEHNLTPVOVQVQSPESFARSHHPKNNKNSLILGLS 992
 QY 961 VDKSVQEPFPHKVVSEHNLTPVOVQVQSPESFARSHHPKNNKNSLILGLS 1020
 DB 961 VDKSVQEPFPHKVVSEHNLTPVOVQVQSPESFARSHHPKNNKNSLILGLS 1020
 QY 993 TGLPDPANPKMLRTCSLPOLSKLFTLMDVPYVGVRODNLEIDIKENIKESPPSSED 1052
 DB 993 TGLPDPANPKMLRTCSLPOLSKLFTLMDVPYVGVRODNLEIDIKENIKESPPSSED 1052
 QY 1021 TGLPDPANPKMLRTCSLPOLSKLFTLMDVPYVGVRODNLEIDIKENIKESPPSSED 1080
 DB 1021 TGLPDPANPKMLRTCSLPOLSKLFTLMDVPYVGVRODNLEIDIKENIKESPPSSED 1080
 QY 1053 IYVEEDTDLOELQASMEQLLRQGEESSEESVYKNSVPEANQTVADDDNPS 1112
 DB 1053 IYVEEDTDLOELQASMEQLLRQGEESSEESVYKNSVPEANQTVADDDNPS 1112
 QY 1081 IYVEEDTDLOELQASMEQLLRQGEESSEESVYKNSVPEANQTVADDDNPS 1140
 DB 1081 IYVEEDTDLOELQASMEQLLRQGEESSEESVYKNSVPEANQTVADDDNPS 1140
 QY 1113 ESALNEEWSHSDSGEIASBCEDSVFNHLEELRLHLEDEMGFEKFEFVYKIAIHDE 1172
 DB 1113 ESALNEEWSHSDSGEIASBCEDSVFNHLEELRLHLEDEMGFEKFEFVYKIAIHDE 1172
 QY 1141 ESALNEEWSHSDSGEIASBCEDSVFNHLEELRLHLEDEMGFEKFEFVYKIAIHDE 1200
 DB 1141 ESALNEEWSHSDSGEIASBCEDSVFNHLEELRLHLEDEMGFEKFEFVYKIAIHDE 1200
 QY 1173 DENIETSKIVONILIGNEHQHLYAKTLHVMADGAYQEDNDE 1214
 DB 1173 DENIETSKIVONILIGNEHQHLYAKTLHVMADGAYQEDNDE 1214
 QY 1201 DENIETSKIVONILIGNEHQHLYAKTLHVMADGAYQEDNDE 1242
 DB 1201 DENIETSKIVONILIGNEHQHLYAKTLHVMADGAYQEDNDE 1242

RESULT 4
 AAU07103 standard; Protein; 1007 AA.
 XX
 AC AAU07103;
 XX
 DT 24-OCT-2001 (first entry)
 XX
 DE Human novel human protein, NHP #3.
 XX
 KW Human; novel human protein; NHP; breast cancer; prostate cancer;
 KW Immunogen; antibody; gene therapy; antisense.
 XX
 OS Homo sapiens.
 XX
 PN WO200161016-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 15-FEB-2001; 2001WO-US05356.

xx Sequence 1035 AA:

Query Match 82.3%; Score 5138; DB 22; Length 1035;
Best Local Similarity 97.3%; Pred. No. 66-302; 0; Indels 28; Gaps 1;
Matches 1007; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

```

OY 208 MKNVITKISGPPVSLHYSYDLSLVSLFRNPRDRSVSLLEKGFIAKRIEFKS 267
DB 1 MKNVITKISGPPVSLHYSYDLSLVSLFRNPRDRSVSLLEKGFIAKRIEFKS 60
OY 268 POLIAEEFCILKTESGSGPIPAKRPASGONISVMPAKITKPAKYGIPLAKKYGDK 327
DB 61 POLIAEEFCILKTESGSGPIPAKRPASGONISVMPAKITKPAKYGIPLAKKYGDK 120
OY 328 KLHEKKPLQKHQAHOQTEPKRVNTGEERRKISEAARRRLEPIEKKOKDOIISLMA 387
DB 121 KLHEKKPLQKHQAHOQTEPKRVNTGEERRKISEAARRRLEPIEKKOKDOIISLMA 180
OY 388 EOMKROEKERLERINRAEQGMWNVLSAGSGEYKAPFLGSGGTIAPSSFSRGOYEYH 447
DB 181 EOMKROEKERLERINRAEQGMWNVLSAGSGEYKAPFLGSGGTIAPSSFSRGOYEYH 240
OY 448 AIFDQMOQRAEDNEAKMKRREIYRGLEPEROKOLAVERAQOYEFLORRREAMONKARA 507
DB 241 AIFDQMOQRAEDNEAKMKRREIYRGLEPEROKOLAVERAQOYEFLORRREAMONKARA 300
OY 508 EGHM-----YILARLQRIORIONENROOKAKIRGEX 539
DB 301 EGHM-----YILARLQRIORIONENROOKAKIRGEX 360
OY 540 KEANHSEQESSEADMRKKITISLKAHANAAVLEKOLEKRRKAYERKKWMEHLV 599
DB 361 KEANHSEQESSEADMRKKITISLKAHANAAVLEKOLEKRRKAYERKKWMEHLV 420
OY 600 AKGVKSSDVSPPLQCHETGSGPSKQOMRSVTSALKEVGVDSITDTRTSEEMOKTN 659
DB 421 AKGVKSSDVSPPLQCHETGSGPSKQOMRSVTSALKEVGVDSITDTRTSEEMOKTN 480
OY 660 NAISKRERILRLNENLKAODEKGMONLSDFEINVEDAKHEKESVSDRRKKWAG 719
DB 481 NAISKRERILRLNENLKAODEKGMONLSDFEINVEDAKHEKESVSDRRKKWAG 540
OY 720 GOLVPIDELTLDTSFSTTEHRTYGEVITKLGPNQSPRAKMGKPTOSVLIIGEAELIQ 779
DB 541 GOLVPIDELTLDTSFSTTEHRTYGEVITKLGPNQSPRAKMGKPTOSVLIIGEAELIQ 600
OY 780 TELLENTTIRSEIPEGEKRYKPLITGEKKYOCISHEINPEAIYDSVVERKSPFESAPQ 839
DB 601 TELLENTTIRSEIPEGEKRYKPLITGEKKYOCISHEINPEAIYDSVVERKSPFESAPQ 660
OY 840 MSIKTEGLPEPDLTEIIOEPSTKNDLSCTITDVAISEKTKTQOSADRTIOE 899
DB 661 MSIKTEGLPEPDLTEIIOEPSTKNDLSCTITDVAISEKTKTQOSADRTIOE 720
OY 900 NEVSEDEVSTVVDLSIDHIEPGTNDOSKCDVDSVOPPEFFKRVHSEHLNLTVOVQ 959
DB 721 NEVSEDEVSTVVDLSIDHIEPGTNDOSKCDVDSVOPPEFFKRVHSEHLNLTVOVQ 780
OY 960 SVOCSPESFAFRSHSLPRKNNKNSLLGLSTGLFDANPMRLTCSLPDLISKLFRL 1019
DB 781 SVOCSPESFAFRSHSLPRKNNKNSLLGLSTGLFDANPMRLTCSLPDLISKLFRL 840
OY 1020 MDVPVGVGDNDNEIDEIKENIKESPSSEDIIVEETDTDQEQASQEQLLRPOPE 1079
DB 841 MDVPVGVGDNDNEIDEIKENIKESPSSEDIIVEETDTDQEQASQEQLLRPOPE 900
OY 1080 EYSEEESEVILKNSVPPRANTDVADDDNPSSESALNEMWSDNSDGTASCECDYAF 1139
DB 901 EYSEEESEVILKNSVPPRANTDVADDDNPSSESALNEMWSDNSDGTASCECDYAF 960
OY 1140 NHLEELRLHLEQEMGEFEFFVEYKIAIHDEDEDNIEICSKIVONILGNHOHLYAKIL 1199

```

```

DB 961 nhleelrlhlegemgfekffeyekikahdedenietcskivnllgnehqnlhyakil 1020
OY 1200 HLYWADGAYGEDNDE 1214
DB 1021 hlywadgaygednde 1035

RESULT 6
AAM40996
ID AAM40996 standard; Protein; 966 AA.
AC AAM40996;
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 5927.
XX
KW Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN M0200153312-A1.
PD 26-JUL-2001.
PE 26-DEC-2000; 2000MO-US34263.
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0532317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Trang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR N-PSDB; AAI60152.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 2; SEQ ID NO 5927; 10078pp; English.
XX
AC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC activation/inhibition activity, chemotactic/chemokine activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX

```

SO	Sequence	966 AA:	
Query Match	75.98;	Score 4735.5;	DB 22; Length 966;
Best Local Similarity	96.38;	Pred. No. 1.1e-277;	
Matches 930;	Conservative 2;	Mismatches 5;	Indels 29; Gaps 2;
OY	278	KTESKFGSOP1-PAKRASGONSTSWPAOKITTKPAKYIPLAVKKYDKKLEHKKPLQ	336
DB	1	khfksfgsalyqlktpasqnsiswmpakltkpaakyiplaykkydkkltkpkpqlq	60
OY	337	KHKQAHOTPEKRYNTGEERKRISEEAARRRIEFIEKEKKOKDQITISLMKEDMKROKE	396
DB	61	khkqhbtpekrvntgeerkliseeaarrrielfiekkkqdkqdlislmkaekmrqke	120
OY	397	RLERINBAROGRNVLASAGSGEVKAPPLGSGGTIASSFSRQGYEHYAIRDOMQOQ	456
DB	121	rlerinareqwrnvlasagsgvkaplfgsggtiapsfsrsgyhyairfdmqgq	180
OY	457	RAEDNEAKMKREIYRGQLPEROKGLAVERAKOVEEFLORKEAKONKARABGHM----	511
DB	181	raedneakmkreiyrglperqgqlaverakoveeflqrkeamqnararegmqlgn	240
OY	512	-----VYLARLROIRLONFNFNROIRKALKEGKREKRNHSGQ	548
DB	241	laamygrpsstrgqprnkeevyarlrlrqlqnerqgkalklqgkkaanhsgq	300
OY	549	EGSEADMRKKITESTLKAHANARAVALKEOLEKRRKEAYERKKVMEHLVAKGVKSSDV	608
DB	301	egseadamrtrkkeslkaahanaaavalkeqlerkeayerkkvwehllvakgvkssdv	360
OY	609	SPLGHEGTGSPSKOQMSVSVTSALKEVGDSSLTDTRFSEMOKTNNATISSKREI	668
DB	361	splghegtgspskqmsvsvtsalkevgdssldttrfseemkktlnaaiskrel	420
OY	669	LRLRLNLKAOEDEKGMQLSDTFEINVEDAKHEHEKESVSSDKKWEAGQVIRPLDE	728
DB	421	lrlrlnenlkaoedekgqlsdtfeinvhedakehekesvsdtkkweagqvlrplde	480
OY	729	LTLDTFSFTEHRTVEEVILKLPNGSPRRMGKSPDVLKILGAEVQLTTELENTTI	788
DB	481	ltdltsfsterhtvgevlklpgnsprrawgkspldsvlkljgaaeiqltellencti	540
OY	789	RSEISPEGEKYKPLITGEKKVOCISHEINPSAIVDSPVETKSPERSASPMOSIKLEGNL	848
DB	541	rseispegekykplitgekkvqistheinpaiavdspetkspetsespmosiklagnl	600
OY	849	EEPDLTETLLOPSTGNDESLPCTITDVTISEEKREKQNSADRITIOENESQCVS	908
DB	601	eeppdltetllpogstgnkndeslpcctitdviseeketkeqsadrlitigenesqcvgs	660
OY	909	STVDLSDIHEPGTNDSDHSCDVNKSVOPEPFHKKVYHSHLVLVQOVSVOCSPES	968
DB	661	stvdqldidhepgtnsdgshscdvksvqpepfthkvyhshlnlvpqvsgvqcspees	720
OY	969	FAFSSHSLPPKNNKNSLLIGSLGLPANNPKMLRTCSLPLDLSKLFRTLMADVPVGDV	1028
DB	721	fafsshslppknnknslligslglpannpkmlrtcslpdlsklrtlmadvplvgdv	780
OY	1029	RODNLDEIKDENIKEGPSDSEDIYFEETDIDLOELASMEQLRQPGGEYSSEESV	1088
DB	781	rqdnldeideidenikegpsdsediyfeetdldlqelqasmeqllrpggeysesv	840
OY	1089	LKNSDVEPTANGTDAVEDDNPSSSALNEEWHSDNSDGETLASECQCSVNHLEERLH	1148
DB	841	lksndveptangtcdaveddnpssesalneewhsdsngdelaseccsdfnhleerlh	900
OY	1149	LEQSMGFKEFEVEYERIKAIHEDEDENIEICKIYONITIGNHGHLYAKITLHVAWAGAY	1208
DB	901	legsmgfekfeyekikaihedenieickivgnllgnehghlyakitlhlvmadgay	960
OY	1209	QEONDE 1214	

DB	961	gednde 966	
RESULT	7		
AA040997			
ID	AA040997	standard; Protein; 966 AA.	
XX			
AC	AA040997;		
XX			
DT	22-OCT-2001	(first entry)	
XX			
DE	Human polypeptide SEQ ID NO 5928.		
XX			
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukaemia.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200153312-A1.		
XX			
FD	26-JUL-2001.		
XX			
PF	26-DEC-2000; 2000WO-US34263.		
XX			
PR	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
PR	29-NOV-2000; 2000US-0727344.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QN, Zhou P, Goodrich R, Drmanac RT;		
XX			
DR	WPI: 2001-442253/47.		
XX			
DR	N-PSDB: AA160153.		
XX			
PT	Novel nucleic acids and polypeptides, useful for treating disorders		
XX	such as central nervous system injuries		
XX			
PS	Example 2; SEQ ID NO 5928; 10078pp; English.		
XX			
CC	The invention relates to human nucleic acids (AA157798-AA161369) and		
CC	the encoded polypeptides (AA038642-AA042213) with nootropic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilisation of the activities such as: Immune system suppression,		
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic		
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukaemias and		
CC	C.N.S. disorders.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification.		
XX			
SO	Sequence	966 AA:	
Query Match	75.98;	Score 4735.5;	DB 22; Length 966;
Best Local Similarity	96.38;	Pred. No. 1.1e-277;	

Matches	930; Conservative	2; Mismatches	5; Indels	29; Gaps	2;
OY	278	KTESKFGSOP1-PAKRPASGONSISVPAOKITTKPAKXGIPALAKKKGDKKLHKRPLO	336		
Db	1	khfskfsgaglyqlkprpasqnsiswmpaqkltkpaakyglpalykkygdckllhckpqlq	60		
OY	337	KHKQAOOTPEKRYNTGEEERKISSEEAARKRRLKEPIEKEKKOKODIISLMKAEOMKROKE	396		
Db	61	khgqahqtpkryntgeertrkiseeaartrrleflekckkygdqllsmkcegmkrkgke	120		
OY	397	RLEKRNARPOGWRNVLSCGSEVYAPLGGCTLAESSFSSRCOYEHYHAIPTQMOQO	456		
Db	121	rletlnareqgwrnvlsgsgsevyapllsgsgtlaessfssrcoyehyhaltdqmqgq	180		
OY	457	RAEDNEAKWRREIYGRQLPEBROKOLAVERAKOVEBFLORRREAMONKARAGHM----	511		
Db	181	raedneakwrrelygrqlperqqlaverakqveeflqrkreanqmkareegllmqn	240		
OY	512	-----VYLARLRQIRLQNFNERQOIKAKLIGEKKEANHSRQ	548		
Db	241	laamygrpsersrgqprkneeeylarlrqlrqlqnfnerqikaklrgkekkeahsegg	300		
OY	549	EGSEADMRKRIEESLKAHANARAVALKEOLRKRREKREKREKWEHLVAKGVSSDV	608		
Db	301	egseeadmrkrileslkahanaraavlkqqlerkrkeayerekkvwehlyvakgykssdv	360		
OY	609	SPPLGHEHTGSGSPSKOOMSIVTSALKEVGVDSLTPTRTSPSMQKTNNATSKREI	668		
Db	361	spplghehtgsgspkqgmsvtsalkevygdsltdtrtseemqktnnalskrei	420		
OY	669	LRLNEMTKROEDKEMONSPPFETNHEDAKHEHKEKSVSDKKKPAAGOLYIPDE	728		
Db	421	lrlnenlkegedekgnlstfelnvheadekeheksevgdqlpide	480		
OY	729	LTLDTSTTERTVGEVITKLPNGSPRANKSPDDSVLKLGAELQLOTELENTT	788		
Db	481	ltdltssttertlvgevlklpngsprankspddsvlklgaelqltelenttll	540		
OY	789	RSEISPEGEKYKPLITGEKKVOCISHEINPSAIVDSPVETKSPERSPOMSLKEGNL	848		
Db	541	rseispegekykplltgekkvgqlshelnpsaivdspevkspetsespsqmslklgnl	600		
OY	849	EEPDDLETELLOEPSCGTNDKESLPCTITDVMISEEKETKETSADRRITQENEVSEDEVS	908		
Db	601	eedpddlelloepsgtndkdeslpctltvmiseeketketsadrritqenevsedegvs	660		
OY	909	STVDQLSDIIEPCTNDKSHSCDVDSVOPRPFHKKVYVSHENLNVPOVOSVQCSPEES	968		
Db	661	stvdqlsdiliepctndshskcdvdksvopprffhkvvhshenlnlvpqvgvqcspees	720		
OY	969	FAFASHHLPKPKNNKNSLLIGLSTGLFPANNPKMLRTCSLPDLAKLFRITLMDVPTVDV	1028		
Db	721	fafashhlpkpknnkslllgstglfpannpkmlrtcsldlklfrltlmdvptvgdv	780		
OY	1029	RONMLTDEIKDENIKESGSDSDYFETDIDLOELQMSQOLLRBOGEXYSEESV	1088		
Db	781	rgonmltdeideenikegspdsdylfetedidloelqmsqollrboqgeyseeesv	840		
OY	1089	LKNSDVPTANGTVADDEDNPSSESLALNEKHNSDNOGELASSECDSVFNHLEELRLH	1148		
Db	841	lknsdvptangtvadeddnpssealaneekhnsdndogelaseecdsavfnhleelrlh	900		
OY	1149	LEQDMGEKFEVEYERIKAIHEDEDENIEGSKITVONTIGNHQLHAKITLHVADGAY	1208		
Db	901	leqdmgektefeyekikaihededenietskivqnllgnehqlhlyaklthlvmadgay	960		
OY	1209	QEDNDE 1214			
Db	961	gednde 966			

RESULT 8
AAV76753

ID	AAV76753 standard; Protein: 345 AA.
XX	
AC	AAV76753;
XX	
DT	17-APR-2000 (first entry)
XX	
DE	Human protein kinase homologue, PKH-6.
XX	
KW	Protein kinase homologue; human; PKH; diagnosis; therapy; cancer; AIDS;
KW	autoimmune disorder; inflammatory disorder; reproductive defect; asthma;
KW	diabetes mellitus; infertility; ovulatory defect; endometriosis;
KW	polycystic ovary syndrome.
OS	Homo sapiens.
XX	
PN	US6013455-A.
XX	
PD	11-JAN-2000.
XX	
PF	15-OCT-1998; 98US-0173581.
XX	
PR	15-OCT-1998; 98US-0173581.
XX	
PA	(INCY-) INCYTE PHARM INC.
PI	Hillman JL, Yue H, Yang YT, Corley NC, Gorgone GA, Azimzai Y;
PI	Lu DAM, Bandman O, Guegler KJ;
XX	
DR	WPI: 2000-136321/12.
XX	
N-PSDB:	AA686797.
PT	Nucleic acids encoding a human protein kinase homolog useful for
PT	preventing, diagnosing and treating cancer, autoimmune/inflammatory
PT	disorders and reproductive defects -
XX	
PS	Claim 1: Column 51-54; 38pp; English.
XX	
CC	This sequence represents a human protein kinase homolog (PKH) of the
CC	invention. The PKH sequences may be used in the prevention, treatment and
CC	diagnosis of diseases associated with inappropriate PKH expression such
CC	as cancers, autoimmune/inflammatory disorders and reproductive defects.
CC	They may be used to treat disorders associated with decreased PKH
CC	expression such as cancers (e.g. lymphoma, melanoma and cancers of the
CC	breast lung and prostate), autoimmune/inflammatory disorders
CC	(e.g. AIDS, asthma and diabetes mellitus), and reproductive
CC	defects (e.g. infertility, ovulatory defects, endometriosis and
CC	polycystic ovary syndrome). The DNA may be administered to treat diseases
CC	by rectifying mutations or deletions in a patient's genome that affect
CC	the activity of PKH by expressing inactive proteins or to supplement the
CC	patients own production of PKH polypeptides. Additionally, the DNA may be
CC	used to produce PKH, according to standard recombinant DNA methodology,
CC	by inserting the nucleic acids into a host cell and culturing the cell to
CC	express the protein. Conversely, antisense nucleic acid molecules may be
CC	administered to down regulate PKH expression by binding with the cells
CC	own PKH genes and preventing their expression by binding with the cells
CC	sequences may also be used as DNA probes in diagnostic assays to detect
CC	and quantitate the presence of similar nucleic acid sequences in samples,
CC	and hence which patients may be in need of restorative therapy. They may
CC	also be used to study the expression and function of PKH polypeptides and
CC	their role in metabolism. The PKH polypeptides may be used as antigens in
CC	the production of antibodies against PKH and in assays to identify
CC	modulators (agonists and antagonists) of PKH expression and activity. The
CC	anti-PKH antibodies and PKH antagonists may also be used to down regulate
CC	PKH expression and activity. The anti-PKH antibodies may also be used as
CC	diagnostic agents for detecting the presence of PKH polypeptides in
CC	samples.
XX	
XX	
Sequence	345 AA:

Query Match 27.8%; Score 1734; DB 21; Length 345;
Best Local Similarity 90.5%; Pred. No. 6,2e-97;
Matches 341; Conservative 3; Mismatches 1; Indels 32; Gaps 1;

```

OY 1 MEKYVRLQKIGESFGKATLVKSTEDGRQYVIKEINISMSSKERESESRREVAVLANMKH 60
DB 1 mekyvrlqkigsgfkgallvkstgedrgyvlkeinsismsskereeerrrevavlanmkh 60
OY 61 PNIVQYRESFENGSLIYVMDYCEGDLFKRINAKGVLFQEDQILDFVQICLAKHVN 120
DB 61 pnivqyresfe-----gllwfvqiclaikhvn 88
OY 121 DRKTLHRDIKSONIFLTRDGTVOLGDFGIARVLNSTVELARCTIGTPYLSPEICEKNPY 180
DB 89 drklhldiksgnifltkdgtvlgdfiglarvlnstvelarctigtpylspeiceknpy 148
OY 181 NNSKDIWALGCVYELCTLKHAFEGAGSMKNLVYKTIISGSPPVSLHYSYDLRSIVSOLF 240
DB 149 nksdiwalgcvlyelctlkhafeagsmknlykltisgsfpvslhysydlrsivsqflk 208
OY 241 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEEFCLTKFSKFSQSPIPAKRPASGQNSI 300
DB 209 rnpdrpsvnsilekglfakrliekflspqlaeeefcltkfskfsqspipakrpasgqnsi 268
OY 301 SVMPAKITTPAAKYGIPLAYKRYGDKLHEKKPLOKHKQAHQTPKRVNTEERRKRISE 360
DB 269 svmpakittpaakygiplaykrygdklhekplqkhaqtpkrvntgeerrkriase 328
OY 361 EAARRRLEFTEKERKQ 377
DB 329 eaarkrlieflekcker 345

RESULT 9
AAE06211
ID AAE06211 standard; Protein: 345 AA.
XX
AC AAE06211;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human protein kinase homolog-6 (PKH-6).
XX
KW Human; protein kinase homolog-6; PKH-6; cytostatic; protein therapy;
KW vaccine; immunosuppressive; antisclerotic; antiabortive; adenocarcinoma;
KW Acquired immune deficiency syndrome; AIDS; melanoma; cancer; bone; liver;
KW breast; autoimmune disorder; multiple sclerosis; drug screening; anaemia;
KW Crohn's disease; ectopic pregnancy; tubal disease; inflammatory disorder;
KW reproductive disorder; polycystic ovary syndrome; asthma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 4..226
FT /note= "Signature sequence"
XX
PN US6264947-B1.
XX
PD 24-JUL-2001.
XX
PE 20-OCT-1999; 99US-0420915.
XX
PR 15-OCT-1998; 98US-0173581.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Bandman O, Tang YT, Hillman JL, Yue H, Guegler KJ, Corley NC;
PI Gorgone GA, Azimzal Y, Lu DAM;
XX
DR MPI: 2001-450728/48.
DR N-PSDB: AAD11848.
XX
PT Human protein kinase proteins and homologs, useful for preventing,
PT diagnosing and treating cancers, autoimmune/inflammatory disorders and
PT reproductive disorders -
XX

```

```

PS Claim 1: Column 51-54; 38bp; English.
XX
CC The present sequence is human protein kinase homolog-6 (PKH-6).
CC Human protein kinase homologs (PKH) and their cDNA molecules are used in
CC the prevention, diagnosis and treatment of diseases associated with
CC increased or decreased expression of PKH. Examples of such disorders
CC include, cancer (e.g. adenocarcinoma, melanoma and bone, breast and
CC liver cancer), autoimmune/inflammatory disorders (e.g. Acquired immune
CC deficiency syndrome (AIDS), anaemia, asthma, Crohn's disease and
CC multiple sclerosis) and reproductive disorders (e.g. tubal disease,
CC ectopic pregnancy) and polycystic ovary syndrome). PKH, its catalytic or
CC immunogenic fragment are used for screening libraries of compounds in any
CC of the drug screening techniques. PKH nucleic acids are used to generate
CC hybridisation probes useful in mapping the naturally occurring genomic
CC sequences. PKH are also used as antigens in the production of antibodies
CC against protein kinases (PK) and in assays to identify modulators of PK
CC expression and activity. PKH is also used in protein therapy.
XX
SQ Sequence 345 AA;
XX
Query Match 27.8%; Score 1734; DB 22; Length 345;
Best Local Similarity 90.5%; Pred. No. 6; 2e-97;
Matches 341; Conservative 3; Mismatches 1; Indels 32; Gaps 1;
OY 1 MEKYVRLQKIGESFGKATLVKSTEDGRQYVIKEINISMSSKERESESRREVAVLANMKH 60
DB 1 mekyvrlqkigsgfkgallvkstgedrgyvlkeinsismsskereeerrrevavlanmkh 60
OY 61 PNIVQYRESFENGSLIYVMDYCEGDLFKRINAKGVLFQEDQILDFVQICLAKHVN 120
DB 61 pnivqyresfe-----gllwfvqiclaikhvn 88
OY 121 DRKTLHRDIKSONIFLTRDGTVOLGDFGIARVLNSTVELARCTIGTPYLSPEICEKNPY 180
DB 89 drklhldiksgnifltkdgtvlgdfiglarvlnstvelarctigtpylspeiceknpy 148
OY 181 NNSKDIWALGCVYELCTLKHAFEGAGSMKNLVYKTIISGSPPVSLHYSYDLRSIVSOLF 240
DB 149 nksdiwalgcvlyelctlkhafeagsmknlykltisgsfpvslhysydlrsivsqflk 208
OY 241 RNPDRPSVNSILEKGFIAKRIEFLSPOLIAEEFCLTKFSKFSQSPIPAKRPASGQNSI 300
DB 209 rnpdrpsvnsilekglfakrliekflspqlaeeefcltkfskfsqspipakrpasgqnsi 268
OY 301 SVMPAKITTPAAKYGIPLAYKRYGDKLHEKKPLOKHKQAHQTPKRVNTEERRKRISE 360
DB 269 svmpakittpaakygiplaykrygdklhekplqkhaqtpkrvntgeerrkriase 328
OY 361 EAARRRLEFTEKERKQ 377
DB 329 eaarkrlieflekcker 345

RESULT 10
AAB95221
ID AAB95221 standard; Protein: 320 AA.
XX
AC AAB95221;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:17342.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PT 28-JUL-2000; 2000EP-0116126.
XX

```

XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INSR.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

XX Claim 8: SEQ ID 17342; 2537bp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises at least 15 nucleotides and the combination
CC of the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB2446 to
CC AAB95993 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 320 AA:

Query Match 22.2%; Score 1387; DB 22; Length 320;
Best Local Similarity 100.0%; Pred. No. 5,2e-76;
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 MNVLTKITSGSPPVSAHSYDLSVQLFRNPRDRPSVSLIEKGFIAKRIEKPIS 267
DB 1 mknlytkitsgsfpvslahsydlrsylsqlifrnprdrpsvsliekfikrilekfs 60
QY 268 POLIAEECLTFSKSGSOPIPAKRPASGONSISVPAOKITTPAKYGIPLAKYKYGK 327
DB 61 poliaeecltfsksgsopipakrpasgonsisvmpaditktpakyyiplaykkygk 120
QY 328 KLEHKKPIQKHQAHOHPERKRVNTGEERRKRISEBAARRRLEFIEKRRKODQIISLMA 387
DB 121 kinehkpriqkhqahqepkrvntgeerrkriiseaarrrlefeiekkqkqdlislma 180
QY 388 EOMKROEKERLERINRAREGQNRNVLSAGSGEVKAPFLSGGCTIPSPSSFGSGVEYEH 447
DB 181 egmkroekerlerinraeregqnrnvlsagsgdevkapflsggctipssfgsgyeyeh 240
QY 448 AIFDPMQOQRAEDNEAKMKREITVGRGLPER 477
DB 241 aifdpmqoqraredneakmkreitygrglper 270

RESULT 11

" , , , "

AAU03545
ID AAU03545 standard; Protein: 649 AA..
XX AC AAU03545;
XX 12-SEP-2001 (first entry)
XX Human protein kinase #45.
XX Human; protein kinase; PK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.

XX Homo sapiens.

XX WO200138503-A2.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US32085.

XX 24-NOV-1999; 99US-0167482.

XX (SUGEN-) SUGEN INC.

XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;

XX WPI: 2001-343950/36.

XX N-PSDB: AAS06745.

XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal associated diseases, and microbial infections -

XX Claim 7: Figure 2; 433pp; English.

XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.

XX Sequence 649 AA:

Query Match 16.4%; Score 1025; DB 22; Length 649;
Best Local Similarity 32.8%; Pred. No. 9.6e-54;
Matches 268; Conservative 125; Mismatches 201; Indels 224; Gaps 25;

QY 1 MEKYVRLOKIGESFKALIVSTEDGRQYVKEINISMSKRESESRREAVLANMKH 60
DB 1 mkyvrvlorkigesfkaliivstedgrqyvkeinismskresesrreavlanmkh 60
QY 61 PNIVQYRESEENGSLEYIMDCGEGDLPFRINACKGVLPQEDDLPDFVQICATLKHV 120
DB 61 pnivqyreseengsleyimdcgegdlpfrinackgvlpqeddlpdfvqicatlkhv 120
QY 121 DRKILHRDRIKSIKRIETKGVY-QLDREGIAVLAVLSVLEARTCGIPYISPEICMKP 179
DB 121 drkilhrdriksikrietkgy-qldregiavlavlsvleartcgyipyispeicmkp 179

```

Db 121 drklhhdkaqnfifskngwakjgdfgiarvlnsmelartcigrpylspcgnkp 180
OY 180 YNKSQDWMALGCVLYELCTLKAHAFEGSKNVLKIIIGSGPPPSLHNSYDLSTVQLP 239
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 ymktcdwslgcvlyelctclhprlqgnmqlvklqahaplspsgrlshlsqglf 240
OY 240 KRNDRPVSNSLLEKGFIAKRIEKLSPQ-----LIAEFLCKTFESK---FGSQPI 288
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 qvsrdpsinsllktpflenllpkylpevsfevltwllaeljgcsrltvtftrcepl 300
OY 289 -PA-----KRPASGQNSISVMPAKITTKPAARYGIPLAYKKYGDKKLHEKPLQKHQAH 342
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 hpacimwfkhnpscflqikmlepkaavcgjdydyaqldmlrrtah--kp-----syh 353
OY 343 QTPKRYVTGEERERKISEEAAARRKRLFEETKEKKOKDOIISLMAEQMKROEKERLERIN 402
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 354 pldge--ntgve-----dygqetrbpsp--sqwpaeylqrfesagqyflk 395
OY 403 RARQGMRNVLASGSGEVPKAPPLGSGGTIAPSSFSRQYBHYHAIFDQMOQORAEONE 462
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 396 vekqlglr-----psaeapn-----ynqrqelrnsngee 423
OY 463 AKMRREITYGRGLPEPROKGLAVERAKOVEEFLQRRKEAMQKARAKGHVYIARLRQRL 522
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 424 prlge--lprfkemkeqeywkgle-----lrrq 450
OY 523 QNFENROQIRAKLKEGKEKEMHSEGGSGSEADMRRKRTESIKAIANRAAVLKEQLERK 582
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 451 qyhdmeleltkmgtrpedl-----eklikmrirgntk-----esk 486
OY 583 RKEAVEREKKWEHEHLVAKGVKSSDVSPPLGQHETGSGSPSKQMSVLSVLSALKEGVND 642
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 487 npe-----qkykakkgvktf-----eind 505
OY 643 SLIDTRETSEMQNTNNAISKRREILRLNENLKAQDEKGMQLSPFE---INVEH 698
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 506 kolsd-----enllqeeamdiptellfedgmktfkye 539
OY 699 DAKEH-----EK-----EKSVS-----DRKKWEAGG-OLVYIPDELTLDTSTFT 738
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 540 cvkehgdyltkafekllcpceagfstqvaavgnrrgwdgagpqltllgmaavadlstcpt 599
OY 739 ERHTVEYIKLGPNGSPFRAMGKSPDTSVKILIGEAEI 776
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 600 gpdn-gvylvleglpgnrtkgwrheapptlmsvlaaahl 636

```

```

PR 19-JUN-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HXSE-) HXSEQ INC.
DR Tang YT, Liu C, Dmanag RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AY, Yang Y, Wejthman T, Goodrich R;
XX WPI: 2001-476283/51.
XX N-PSDB; AAK51477.
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PS useful in diagnosis and gene therapy -
XX
XX Claim 20; Page 3232-3233; 6221pp; English.
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX Sequence 506 AA:
SQ

```

```

Query Match 14.18; Score 882; DB 22; Length 506;
Best Local Similarity 42.58; Pred. No. 3e-45;
Matches 175; Conservative 70; Mismatches 103; Indels 64; Gaps 6;
OY 1 MEKYLRLQKIGESFGKALIVKSTEDGRQYVKEINISMSSKERESREYAVLIANKH 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 mddymvlrmigsgfgrtallvghessnqmfaekrlprk-sfsntqnsrkeavvllakmh 59
OY 61 PNIVQRESFEENGSLVYMDYCEGDFRINAQGVYFOEDQIIDWFVQICLAKHVN 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 60 pntvafkesfeeghlylmeycdgdqldmqkirkqgkllfpedmlnwftqmclyvnhlh 119
OY 121 DRKILHROKSONTFLTFDGVOLGDFGIARVNSTVELARPCIGPYLSPSEICENRPY 180
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 krrvlhrdksknlltqngkvklgdfgsarllsnpmafaclyvgpyvyppeilwenipy 179
OY 181 NNRKSDIMALGCVLYELCTLKAHAFEGSKNVLKIIIGSGPPPSLHNSYDLSTVQLP 240
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 mksdswslgcllyelctclhprlqgnmqlvklqahaplspsgrlshlsqglf 239
OY 241 KRNDRPVSNSLLEKGFIAKRIEKLSPQLIAEFLCKTFESKSGSPRIKAPKASGQNSI 300
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 rnpshpsactllstrgavrtvqkclpeltme----- 272
OY 301 SYMPAOKITTKPAKYGIPLAYKKYGDKKLHEKPLQKHQAHQOTPEKRYNTEERKISE 360
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 273 -----ygeevleelk-----nshntcpkknupstiritalgn 304
OY 361 EAARRRLEETKEKKOKDOIISLMAEQMKROEKERLERINRARGQMRNV 412
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 eastvg-----eeqdrkgshtdlesine--nlvesalltrvr-eeqgnksv 348

```

RESULT 13
AAM79328

ID AAM79328 standard; Protein: 527 AA.
 AC AAM79328;
 DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 2974.
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 XX tissue growth factor; immunomodulatory; cancer; leukaemia;
 XX nervous system disorder; arthritis; inflammation.
 OS Homo sapiens.
 PN WO200157190-A2.
 PD 09-AUG-2001.
 XX 05-FEB-2001; 2001MO-US04098.
 PF 03-FEB-2000; 2000US-046914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620323.
 PR 01-SEP-2000; 2000US-0654561.
 PR 15-SEP-2000; 2000US-0653561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HXXE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,
 PI Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW,
 PI Xue A, Yang Y, Wehrman T, Goodrich R.
 XX WPI: 2001-476283/51.
 DR N-PSDB: AAK52461.
 DR Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX Claim 20: Page 218; 6221pp; English.
 PS The invention relates to polynucleotides (AAK51456-AAK51435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX Sequence 527 AA:

Query Match 14.0%; Score 875; DB 22; Length 527;
 Best Local Similarity 42.2%; Pred. No. 8.4e-45;
 Matches 174; Conservative 70; Mismatches 104; Indels 64; Gaps 6;

QY 1 MEKVRLOKIGSGFKALIVKSTEDGROYVIREINSRMSKRESESEVAVLANMKH 60
 DB 22 mdyymvlmigeqsfrelllyghessngmfamkelrlpk-sfentqnsrkeavllakmh 80
 QY 61 ENIVYXRESENGSLTYVMDYCGSDLPKRINAGGVLPQEQIILDMVQCLAKNHN 120
 DB 81 pvlvafkesfaeaghllylvmeycdggdlmkikqgkylfpedmllnwftqmclyvnhn 140

QY 121 DRKILHRDIKSONIFLTQDGTVOJGDFGIARVLNSTVELARFCIGTPYVLSPECENKPY 180
 DB 141 kkrvlnrdlksknllftngkglgdfgsarlspmafactyvgtyyvpelwani 200
 QY 181 NNRSDIWAIGCVLYELCTLKAHFAEGSMKNLVLKIIISGSPVSLHYSYDLRSLSQLFK 240
 DB 201 nksdlsyglcllyelctlkhpfganswnllkvcgcslpsbhsyqlfvlxqmfk 260
 QY 241 RNRDRPSVNSILKFKRIKREFLSPOLAEFECLTFKFCGSPILPAKRPASGONSI 300
 DB 261 rnpshpsatlslrglvarlvqclpelpme----- 293
 QY 301 SVMPAQKITRPAKAVGICPLAYKKYGDKKLHEKKPLOKHAQHPKRVNTEGERRKISE 360
 DB 294 -----ygeevleelk-----nshhnprrkktppsrllalgn 325
 QY 361 EAARKRLERIEKKOKDQIISLMAKAGOMKROEKERLERINRAREGGRNV 412
 DB 326 eastvg-----eeqdkshdlesine--nlvesalrrvr-eekgnksv 369
 RESULT 14
 ID AAY68778 standard; Protein: 510 AA.
 AC AAY68778;
 XX 16-MAY-2000 (first entry)
 DE Amino acid sequence of a human phosphorylation effector PHSP-10.
 XX Human; phosphorylation effector; PHSP; proliferative disorder;
 XX immune disorder; neuronal disorder.
 XX Homo sapiens.
 OS
 PA Key
 PI Modified-site 24 Location/Qualifiers
 PI Modified-site 51 /note="potential phosphorylation site"
 PI Modified-site 51 /note="potential phosphorylation site"
 PI Region 52..261
 PI Modified-site 152 /note="protein kinase family signature sequence"
 PI Modified-site 185 /note="potential phosphorylation site"
 PI Modified-site 201 /note="potential glycosylation site"
 PI Modified-site 210 /note="potential phosphorylation site"
 PI Modified-site 247 /note="potential phosphorylation site"
 PI Modified-site 292 /note="potential phosphorylation site"
 PI Modified-site 297 /note="potential phosphorylation site"
 PI Modified-site 312 /note="potential phosphorylation site"
 PI Modified-site 323 /note="potential phosphorylation site"
 PI Modified-site 325 /note="potential phosphorylation site"
 PI Modified-site 329 /note="potential phosphorylation site"
 PI Modified-site 349 /note="potential phosphorylation site"
 PI Modified-site 358 /note="potential glycosylation site"
 PI Modified-site 377 /note="potential phosphorylation site"
 PI Modified-site 381 /note="potential phosphorylation site"

Wed May 15 14:13:54 2002

us-09-783-320-4.rag

Page 14

Sequence 698 AA;

Search completed: May 15, 2002, 07:52:53
Job time: 4931 sec

Query Match 10.1% Score 628; DB 22; Length 698;
Best Local Similarity 23.8% Pred No. 1e-29;
Matches 227; Conservative 116; Mismatches 271; Indels 340; Gaps 29;

```
QY 1 MEKYYRLQKIGESGKALIVKSTEDGQVYKEINISRMSSKREBSREAVLANMKH 60
DB 1 MEKYYRLVGVGATVHICIKADQKIVIKQIVGQMKKEERQAQECQVLIKILNH 60
QY 61 PNIVQRESFEENGSLYVMDYCEGDLFKRINQKGVLFQEDQILDMFVQICLAKHVA 120
DB 61 PNIVLEYENFLEDKALAMEYAPGTLAEFLKRCNSLLEETLHFVQILALHVAH 120
QY 121 DRKILARDIKSONIFLRKDG-TVQLDGPGIARVLNSTVELARTCIGTPYVLSPEICENRP 179
DB 121 THLLHRLDKTQNLIDKHMVVKIGDFGSKLSK-----SLPCYISPDLCEGKP 172
QY 180 YNNKSDIWLGCVYELETCTKHAFAEGSMKNLYLKISGFPVSLHYSDRLSVQLF 239
DB 173 YNQKSDIWLGCVYELETCTKHAFAEGSMKNLYLKISGFPVSLHYSDRLSVQLF 232
QY 240 KNPDRPVSNTLEKGFIAKRIEKLSPOLAEFLCTFKFSKFGSQPIPAKRPASGQNS 299
DB 233 SLPEAGRPPLSHM-----AGPLCIRAL----- 255
QY 300 ISVMPAKITKPAKYGIPLAKYKGDKLIHE-----KKPLQKHQAHOPEKRVNTGEE 354
DB 256 -----LHLDGREVVPQINREQDNGCPDQ----- 281
QY 355 RKKISEBARKKRLEFIEKEKKOKDQIISLMAKABQMKROEKERLERINRAREQGMNVLS 414
DB 282 -----RGLIM 286
QY 415 AGGSEVKAAPLGGGTAPSSFGQYEHYNAIFDOMQOQRAED----- 460
DB 287 LFQS-----GNGCLGHSGLDLSQPTLVEALIGYEMVQVGEALSFLLGSAPLDQ 338
QY 461 -----NEAKMKREIYGRG-----LPERQKQL-AVERAKOVEEF--LQRRK 498
DB 339 EPLISDIGHASAAVTGEADIGSGDVNRIPSWERGHLLAGVASSTCVSTFSEGDCKEPD 398
QY 499 EAMQKARAREGHMVY-----LARLRQIRLQNFNRQOTKAKLGEKKEANHSEG-- 547
DB 399 KECWRHKGTGHLIYFASDCVRHSLHSHVHNCNSR-----LKSSSEDSSSRGAG 452
QY 548 -----OEGSEFADMRRKKTESLKAHANARAVALKEQL-----ERRKKE 585
DB 453 PLSCHVLEAPCELTPEEENHVERIFRYGCKSYRPSVAVALHPLHYHEGAGDILNKKKTR 512
QY 586 AVEREKKWEHLVAKYKSSDVAPPLGQHFETGSPSKQOMRSYISVTSALKLEVYDSSL 645
DB 513 RRRKSKPP-----LPCGVGPAASPDIGTSMATGTP-----DSTLA 547
QY 646 TDPRETSEMOXTNNAISSKREILRLNENLKAQEDKGMONLSDFEINVEDAKREHK 705
DB 548 PLIIVRSE--SPTKGQSK--VIKKYKKKKKKKKKEKDEKDEMD-----EKAKIKKK 592
QY 706 EKSVSSDRKKWAGQLVPLDELTLDTSFSTRHTVGEVIRKLPNGSPRRAMGKSPD 765
DB 593 AKKGILTKKK-----SPVKLEP-----SPPD 613
QY 766 SVLKILGEAELOLOTELENTTIRSEISPEGEKYKPLITGEKKVOCISHEINPSALVDP 825
DB 614 -VSRLSARQIARME--SSPESRELESE-DSY--NGRGGLSSEDI----- 656
QY 826 VETKSPERSE-----ASPQMSIKIEGSLLEPPDLETEIILQEPGCTNKDES 870
DB 657 VESSSPKRENTVGAQKGTGAKPSGAKRYNKR-----KSPGSGNPGLS 698
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (COPY)